AD_____

Award Number: DAMD17-03-1-0134

TITLE: Automatic Organ Localization for Adaptive Radiation Therapy for Prostate Cancer

PRINCIPAL INVESTIGATOR: Sarang C. Joshi, D.Sc.
                            Mark Foskey
                            Brad Davis

CONTRACTING ORGANIZATION: University of North Carolina at Chapel Hill
                          Chapel Hill, NC 27599-4100

REPORT DATE: May 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
              Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
                        Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20060207 009

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-05-2005	Annual	05/01/04-04/30/05

4. TITLE AND SUBTITLE
Automatic Organ Localization for Adaptive Radiation Therapy for Prostate Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-03-1-0134

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Sarang C. Joshi, D.Sc.
Mark Foskey
Brad Davis

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-4100

sarang_joshi@med.unc.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT: Abstract on next page.

15. SUBJECT TERMS
Adaptive Radiation Therapy, Conformal Radiation Therapy, Automatic Image Segmentation, Computer-assisted Diagnosis, Digital Imaging

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	45	19b. TELEPHONE NUMBER *(include area code)* 301-619-7325

Abstract

Breast cancer incidence and outcomes vary in women of different racial/ethnic backgrounds. Race/ethnicity and tumor biology may affect outcomes. Since regional lymph node status and tumor markers are strong prognostic indicators, this study examines the role of sentinel lymph node status (SLNS) and cyclin E levels in outcomes for women of various races/ethnicities with breast cancer. Data was collected for 400 women from two cohort groups using existing database and medical records. Data included tumor size, nodal status, estrogen receptor status, HER-2/neu status, cyclin E levels and race/ethnicity. A new database organizes unique study data: socioeconomic status and health-related behaviors. Data quality checks and abstraction continue. Subjects will be matched for as many factors as possible. The final sample of 50 Whites/non-Hispanic and 50 others, including Hispanics, will be analyzed to correlate SLNS to race/ethnicity, cyclin E levels to race/ethnicity and SLNS to cyclin E levels. Disease-free survival and overall survival rates cannot be determined for several years and thus are not available during the award period. It is hypothesized prognostic accuracy of SLNS and cyclin E levels are independent of racial/ethnic factors. This finding would suggest SLNS and cyclin E levels could discriminate outcomes within different racial/ethnic groups.

Table of Contents

Introduction

The goal of this research is to develop automated methods to analyze daily CT scans taken during prostate therapy, in order to make adaptive radiation therapy (ART) more effective and readily implemented. The central technique is deformable image registration, which allows a CT image acquired for treatment planning to be deformed to match each daily image. The resulting deformations can then be applied to contours drawn at planning time, to generate segmentations of treatment images. Inverse deformations can also be applied to dose distributions (represented as images indicating the dose delivered to each voxel). In this way, the dose distributions can be deformed into the reference frame of the planning image, and summed over the course of treatment to determine the total delivered dose. Over the past year we have demonstrated both of these techniques on data from multiple subjects.

Body

Validation of Automatic Segmentation

We performed a statistical validation of our algorithm by performing segmentation on 48 treatment images from four patients. We found that automatic segmentations derived (by deformation) from one individual's planning contours matched that individual's later segmentations at least as well as did segmentations by a different individual. Our analysis was based on two measures of geometric variation, centroid difference and volume overlap. Extensive details of this work are given in two related papers, included as Appendices A and B. With this work, all portions of task 1 in the Statement of Work have been addressed: The algorithms accommodate rectal filling, dealing in particular with the severe problems caused by gas (1a). We have developed and applied methodologies to validate the algorithms (1b). The performance of the algorithm has been analyzed using data sets provided by MSKCC as well as locally acquired data (1c). And our validations quantify inter-rater variability (1d). There remains work to strengthen our results as our patient sample grows.

Dose Accumulation

The ability to transform the distribution of dose from one treatment day to the planning day, and then accumulate such doses from different days, makes it possible to analyze how the motion of an actual patient would affect a treatment plan for that patient. It is thus possible to test novel treatments against motion data from patients that have already been treated, in what we call a virtual clinical trial.

When delivered dose is accumulated, the doses delivered to each voxel can simply be summed over the range of treatments. However, it is also possible to apply models that take into account the effect of fraction size on the biological effectiveness of a dose. We have used the linear-quadratic model, as described in Appendix B.

As a baseline, and to illustrate this concept of dose accumulation, for nine subjects we have evaluated the total delivered dose to the prostate and rectum assuming that no adaptive radiation therapy had been used. The results are shown in Figure 1, and will be presented as a poster at the October meeting of the American Society of Therapeutic Radiology and Oncology (ASTRO). For the prostate, we computed prescribed and delivered EUD, and delivered mean dose. For the rectum, we computed the percent of the volume receiving more than 65 Gy, and also the percent receiving more than 40 Gy. These values can be compared to prescribed values

of 78 for the dose, no more than 17% of the rectum to receive more than 65 Gy, and no more than 35% to receive more than 40 Gy.

| Patient	Prescribed prostate EUD (Gy)	Delivered prostate EUD	Delivered prostate mean dose (Gy)	% of rectum > 65 Gy	% of rectum > 40Gy
1	77.9	74.7	76.3	2	24
2	76.6	73.9	75.6		11
3	78.6	77.8	78.0	1	15
4	75.6	75.9	76.2	9	33
5	76.2	75.3	75.9		8
6	76.8	76.8	76.8	21	43
7	76.0	73.5	74.4		8
8	76.0	77.9	78.0	3	13
9	77.3	75.6	75.7	32	81

Table 1. Prescribed and delivered prostate EUD, delivered prostate mean dose, and the percent of the rectum receiving greater dose than the thresholds of 65 Gy and 40 Gy.

As an illustration of the concept of a virtual clinical trial, for three patients we have computed DVHs of delivered BED using the linear-quadratic model. We used α/β values of 1.5 for the prostate and 4 for the rectum [Brenner 2003]. The patients had been treated with a conventional course of 39 fractions at 2 Gy per fraction. For comparison, we simulated a hypofractionation plan with 15 fractions of 3 Gy each. Using the deformation fields computed for the patients, we computed delivered BED for 15 fractions (using the first 15 imaged treatments for each patient), and compared it to the planned BED for the hypothetical hypofractionation plan, along with the planned BED for the plan actually used. The results are shown in Figure 1.

This work addresses task 2 and provides a framework for task 3.

Key Research Accomplishments

- Comparison of manual segmentation with our automatic method, using several measures, indicating that automatic segmentations derived from one individual's planning segmentation match that person's later segmentations at least as well as manual segmentations by a different rater.
- Computation of the actual cumulative dose delivered to both the cancerous and critical healthy tissues of nine subjects.
- Dose volume histograms of BED using the linear-quadratic model for three subjects.
- Article submitted to *Physics in Medicine and Biology*: "Large deformation 3D image registration in image-guided radiation therapy," by Mark Foskey, Brad Davis, Lav Goyal, Sha Chang, Ed Chaney, Nathalie Strehl, Sandrine Tomei, Julian Rosenman, and Sarang Joshi. Included here as Appendix B.
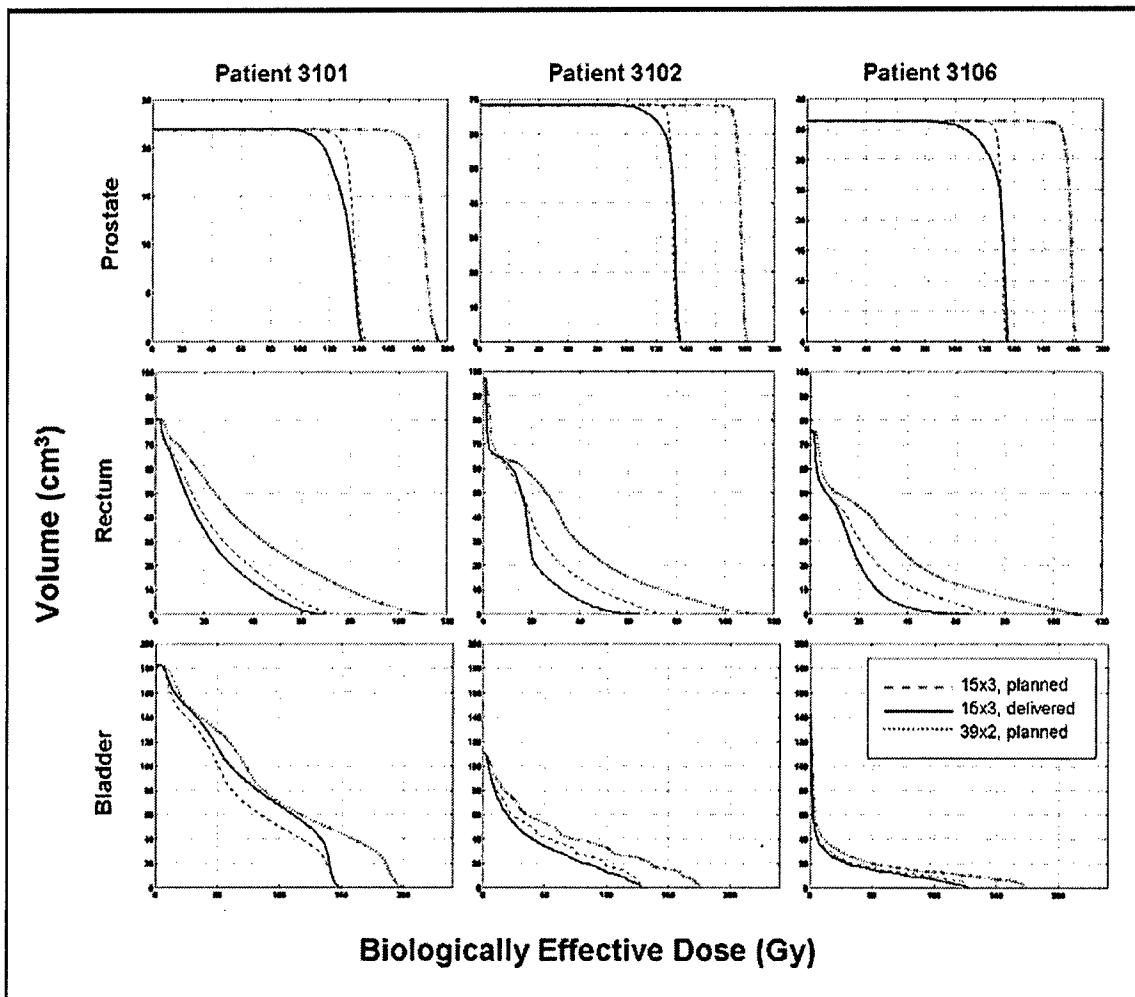
Figure 1. BED for prostate, bladder and rectum. Planned BED is compared to calculated delivered BED for a treatment of 15 fractions at 3 Gy/f, and also to the planned BED for a conventional 39 x 2 plan.

Reportable Outcomes

Abstracts:

M. Foskey, B. Davis, L. Goyal, S. Chang, J. Rosenman, and S. Joshi. "Calculating biological effective dose in the presence of organ deformation." AAPM 47th Annual Meeting, 2005.

K. Wijesooriya, V. Dill, L. Dong, R. Mohan, S. Joshi, E. Weiss, and P. Keall. "Comparison of auto contouring with manual contouring: A first step towards automated 4D treatment planning." AAPM 47th Annual Meeting, 2005.

M. Foskey, B. Davis, L. Goyal, S. Chang, J. Rosenman, and S. Joshi. "Automatic contouring via deformable image registration." ASTRO (Accepted), 2005.

Full Length Conference Papers:

B. Davis, M. Foskey, J. Rosenman, L. Goyal, S. Chang, S. Joshi. "Automatic segmentation of intra-treatment CT images for adaptive radiation therapy of the prostate." To appear in *Proceedings of MICCAI*, 2005. Included here as appendix A.

Conclusions

In the past year we have statistically validated our segmentation algorithm, showing its variation to be within the range of human variation. We have also shown how dose accumulation algorithms based on image registration can be used to analyze the performance of various treatment plans in the presence of organ motion. In the remainder of the grant period we hope to extend these results and take steps toward making them more usable in the clinic.

References

[Brenner 2003] D. Brenner. Measurement of patient positioning errors in three dimensional conformal radiotherapy of the prostate. Int J Radiat Oncol Biol Phys 57:912-914, 2003.

Appendices

Appendices A and B follow.

7

APPENDIX A

Automatic Segmentation of Intra-Treatment CT Images for Adaptive Radiation Therapy of the Prostate

B.C. Davis[1,2], M. Foskey[2], J. Rosenman[2], L. Goyal[2], S. Chang[2], and S. Joshi[1,2]

[1] Department of Computer Science, University of North Carolina, USA,
{davisb,joshi}@cs.unc.edu,
[2] Department of Radiation Oncology, University of North Carolina, USA.

Abstract. We have been developing an approach for automatically quantifying organ motion for adaptive radiation therapy of the prostate. Our approach is based on deformable image registration, which makes it possible to establish a correspondence between points in images taken on different days. This correspondence can be used to study organ motion and to accumulate inter-fraction dose. In prostate images, however, the presence of bowel gas can cause significant correspondence errors. To account for this problem, we have developed a novel method that combines large deformation image registration with a bowel gas segmentation and deflation algorithm. In this paper, we describe our approach and present a study of its accuracy for adaptive radiation therapy of the prostate.

1 Introduction

One major treatment method for prostate cancer is external beam radiation therapy, which uses high energy x-rays that are delivered in a series of 40 or more daily treatments. To be safe and effective, the radiation dose to the cancer-containing prostate should be as high as possible while the dose to surrounding organs such as the rectum and bladder must be limited. This effect is achieved by using multiple radiation beams that overlap on the tumor and are shaped to exclude normal tissue as much as possible. However, internal organ motion and patient setup errors present a serious challenge to this approach. The prostate, rectum, bladder and other organs move in essentially unpredictable ways, and even small changes in their position can result in either tumor under-dosing, normal tissue over-dosing, or both.

Adaptive radiation therapy (ART), which uses periodic intra-treatment CT images for localization of the tumor and radiosensitive normal structures, is being investigated to meet this challenge. In this method a feedback control strategy [1] is used to correct for differences in the planned and delivered dose distributions due to spatial changes in the treatment volume early in the treatment period.

Although in-treatment-room CT scanners provide the enabling imaging hardware to implement ART, no software methods or tools for automatic image processing exist to enable the incorporation of these images in the adaptive

treatment of prostate or other cancer. As a result, all such work must be done manually. However, manual segmentation of the tumor and neighboring organs places an impractical burden on highly skilled and already overburdened personnel. Moreover, clinically significant inter- and intra-user variability of manual segmentations introduces a source of treatment uncertainty that current adaptive radiation therapy techniques do not address [2, 3].

We have been developing an approach for automatically quantifying organ motion over the course of treatment. Our approach is based on deformable image registration, which makes it possible to establish a correspondence between points in images taken on different days. This correspondence can be used to study organ motion and to accumulate inter-fraction dose. In prostate images, however, the presence of bowel gas can cause significant correspondence errors as no correspondence exists for pockets of gas across different days. Shown in Figure 1 are two rigidly aligned axial images of a patient taken on two different days. Due to the transient nature of bowel gas, it is present in one of the days but absent in the other. To account for this problem, we have developed a novel method that combines large deformation image registration with a bowel gas segmentation and deflation algorithm. In this paper, we describe our approach and present a study of its accuracy for adaptive radiation therapy of the prostate.
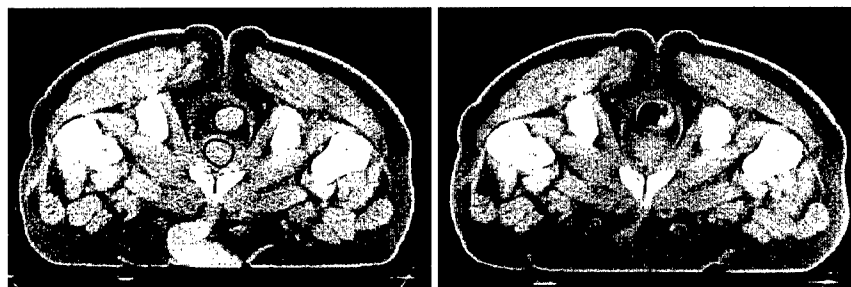

Fig. 1. Axial CT slices of the same patient acquired on different days, showing the effect of bowel gas.

2 Methods

We use the CT taken at planning time, the *planning image*, as a reference. On each treatment day, the patient is positioned and then, prior to treatment, a new CT scan is acquired using an in-treatment-room CT scanner that shares a table with the linear accelerator (Figure 2). Each *treatment image* characterizes the patient configuration at that treatment time.

If there were absolutely no organ motion then the planning and treatment images should all be the same, except for noise from the imaging device. However, because there is organ motion, these images will differ, and the difference
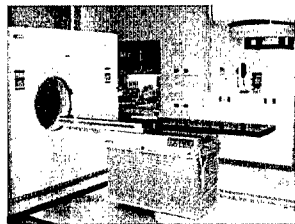
Fig. 2. A treatment room set up for ART: the linear accelerator shares a table with a CT scanner. Daily treatment images are used to track organ motion over the course of a multi-week treatment.

characterizes the organ motion. We have understood the motion when we can tell, for each point in the planning image, which point in the treatment image it corresponds to. In this way organ motion and image registration are linked—we can understand organ motion if we can estimate image correspondence.

We can view an image as a function I from the spatial domain $\Omega \subset \mathbb{R}^3$ to an intensity value in $\mathbb{R}$. Image correspondence is expressed as a function $h \colon \Omega \to \Omega$, called a *deformation field*. For $x \in \Omega$, $h(x)$ is the point in the treatment image, I_T, that corresponds to the point x in the planning image, I_P.

The transformation h is estimated as follows. First, the planning and treatment CT data sets are rigidly registered. This quantifies the rigid patient setup error. In order to accommodate bowel gas we apply our algorithm for segmenting and deflating bowel gas to produce deflated images I_{P_d} and I_{T_d}. Finally, I_{P_d} and I_{T_d} are registered using a high dimensional large-deformation image registration algorithm. h is defined as the composition of these transformations.

Rigid Registration The planning and treatment images are thresholded so that only bone is visible. The region of interest is restricted to the pelvis as it remains fixed while the femurs and spine can rotate or bend. The rigid transformation, r, is estimated by an intensity based gradient descent algorithm [4].

Accommodating Bowel Gas As the contrast between gas and surrounding tissue is very high in CT images, we create a binary segmentation of the gas in an image using a simple thresholding operation. We refine this binary segmentation using a morphological open operation. Next, we construct a deflation transformation s based on a flow induced by the gradient of the binary image. Points along the gas-tissue border, where the gradient is non-zero, flow in the direction of the gradient. As a result, gas filled regions collapse toward their medial skeletons—deflating like a balloon.

We construct a non-diffeomorphic deflation transformation $s \colon \Omega \to \Omega$ such that $I(s(x))$ is the image $I(x)$ after a deformation that deflates gas. The transformation s is constructed by integrating velocity fields $v(x, t)$ forward in time, i.e. $s(x) = x + \int_0^1 v(s(x, t), t)\, dt$. These velocity fields are induced by a force function $F(x, t) = \nabla(I \circ s_t)(x)$ that is the gradient of the binary image. The force function and velocity fields are related by the modified Navier-Stokes operator $(\alpha \nabla^2 + \beta \nabla (\nabla \cdot) + \gamma)v(x, t) = F(x, t)$. We solve for s using an iterative greedy method.

Figure 3 shows the result of our gas deflation algorithm. The large pocket of gas present in the image has been deflated, resulting in an image that can be accurately registered using deformable image registration.
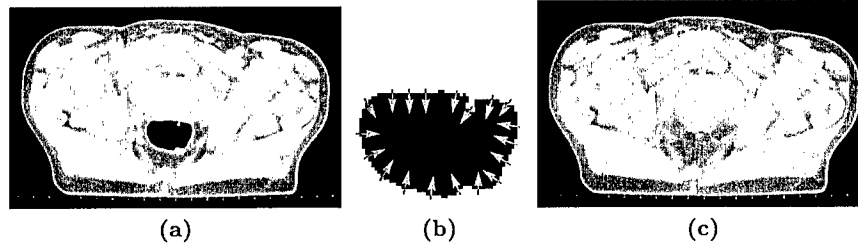
Fig. 3. Gas Deflation Algorithm. (a) CT image with considerable bowel gas. (b) Zoomed in on the gas pocket. The gas is segmented using simple thresholding. Gas is deflated by a flow induced by the gradient of the binary image. (c) The deflated image.

Deformable Image Registration We apply the theory of large deformation diffeomorphisms [5–7] to generate a deformation $h_{def}: \Omega_{P_d} \rightarrow \Omega_{T_d}$ that defines a voxel to voxel correspondence between the two gas deflated images I_{P_d} and I_{T_d}. The registration is determined by finding the deformation field h_{def} that minimizes the mean squared error between I_{P_d} and the deformed image $I_{T_d} \circ h_{def}$,

$$D(h) = \int_{x \in \Omega} |I_{P_d}(x) - I_{T_d}(h_{def}(x))|^2 \, dx.$$

Following citechristensen96,miller99 the transformation is constrained to be diffeomorphic by enforcing that it satisfy laws of continuum mechanics derived from visco-elastic fluid modeling.

Composite Transformation Correspondence between the original images I_P and I_T is estimated by concatenating the rigid, deflation, and deformable registration transformations, i.e.

$$h_{P \rightarrow T} = r(s_T(h_{def}(s_P^{-1}(x))))).$$

This composite transformation is not guaranteed to be diffeomorphic. However, the non-diffeomorphic part of the transformation is restricted to the region of the rectum that contains gas—where no correspondence exists.

Figure 4 shows the result of the method described above. Panel (b) shows the result of automatic segmentation using only large deformation image registration. Manually drawn contours of the prostate and rectum are mapped, using this correspondence, from the reference image (a) onto the daily image. Manual contours are drawn in red while mapped contours are drawn in yellow. Notice the misalignment of the manual and automatically generated contours in the daily image; the presence of bowel gas has caused correspondence errors around the

rectum. A more accurate correspondence between the reference and daily images is established by concatenating registration and deflation transformations as shown in panel (c). Notice the close alignment between the manual contours and the contours generated by our method.

(a) (b) (c)
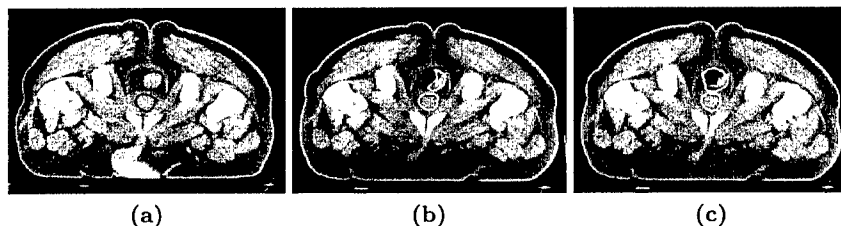
Fig. 4. Automatic segmentation of the prostate and rectum. Manually segmented structures in the planning image (a) are mapped to the daily image (b) *before* accounting for bowel gas, and (c) *after* accounting for bowel gas with our gas deflation algorithm. Manually drawn contours are shown in red and mapped contours are shown in yellow.

3 Results

We now present detailed statistical analysis of the application of our methods to a set of 40 CT images from 3 patients undergoing ART in our clinic. Each CT scan was collected on a Siemens Primatom CT-on-rails scanner with resolution 0.098x0.098x0.3 cm. We analyze the accuracy of our method by comparing automatically generated segmentations to manual, hand-drawn, segmentations. Because of inter-rater variability, however, there is no ground truth manual segmentation to compare against. We therefore compare our automatically generated segmentations with the segmentations from two different manual raters, and then make the same comparisons between the segmentations from the manual raters.

The experimental setup is as follows. The planning image for each patient is manual segmented by rater A. Each treatment image is manually segmented twice, once by rater A and once by rater B. For each patient, our method is used to compute the transformations h_i that map the planning image onto the treatment image for each day of treatment i. An automatic segmentation is generated for each treatment image by applying h_i to the segmentation in the planning image. We can consider our automatic method for producing segmentations as rater C (for "computer").

Each segmentation is represented by a triangulated surface. For manual segmentations, the surface is constructed by applying the power crust algorithm [8] to a set of contours drawn in the axial plane by the manual raters. For automatic segmentations, the surface is generated by applying a transformation h to the vertices of the surface given by the manual segmentation in the planning image.

For each patient and for each treatment day, we make three comparisons: **CA**, automatic segmentation verses manual segmentation by rater A; **CB**, automatic segmentation verses manual segmentation by rater B; and **BA**, manual segmentation by rater B verses manual segmentation by rater A. It should be emphasized that the automatic segmentations are produced by transforming manual planning segmentations produced by rater A, not rater B. Thus, we expect the CA comparisons to be more favorable than the CB comparisons.

In the rest of this section, we present the results of this experiment when comparing centroid differences and relative volume overlap of segmentations.

Centroid Analysis The centroid of the prostate is especially important for radiation treatment planning and therapy because it is the origin, or isocenter, for the treatment plan. To measure the accuracy of our automatic segmentations with respect to centroid measurement, we compare the centroid of each automatic segmentation with the centroids of the corresponding manual segmentations. The differences in the lateral (X), anterior-posterior(Y), and superior-inferior (Z) directions are measured separately.

Figure 5 shows box and whisker plots of these differences for CA, CB, and BA comparisons. All measurements are made in centimeters. Additional summary statistics are presented in table 1.
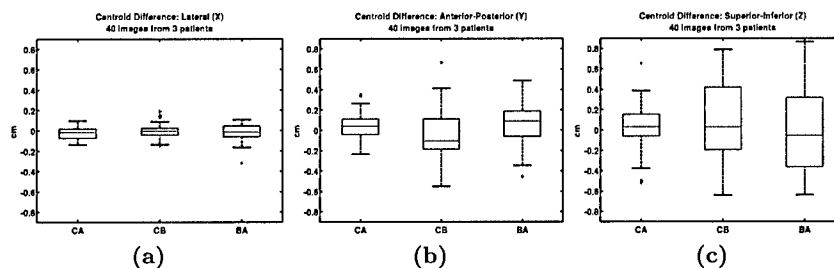
Fig. 5. Centroid differences in the lateral (X), anterior-posterior (Y), and superior-inferior (Z) directions (cm). The horizontal lines on the box plots represent the lower quartile, median, and upper quartile values. The whiskers show the extent of the rest of the data. Outliers, which fall outside 1.5 times the interquartile range, are denoted with the '+' symbol.

Shown in Table 1 are the 99% confidence intervals for the true mean of each distribution of centroid differences. The confidence intervals for the means of the CA and CB differences both overlap with the confidence interval of the differences between human raters (AB), and are on the order of one voxel. Note that the superior-inferior (Z) direction has a slice thickness of 0.3 cm. We conclude that the automatic segmentation method is as accurate for estimating centroids as human raters and, as seen by the standard deviations, just as reliable.

Relative Volume Overlap Analysis A measure often reported for comparison of segmentations is relative volume overlap. This measure has been defined

Centroid Difference Summary (cm)								
	Lateral (X)			A-P (Y)			Sup-Inf (Z)		
	CA	CB	BA	CA	CB	BA	CA	CB	BA
mean	-0.026	-0.007	-0.022	0.035	-0.052	0.070	0.022	0.065	-0.046
median	-0.018	-0.004	-0.015	0.040	-0.104	0.089	0.030	0.028	-0.054
std. dev.	0.06	0.07	0.08	0.14	0.23	0.20	0.24	0.38	0.38
99% CI min	-0.047	-0.030	-0.049	-0.010	-0.129	0.007	-0.054	-0.058	-0.167
99% CI max	-0.006	0.016	0.004	0.081	0.023	0.133	0.10	0.189	0.073

Table 1. Summary statistics showing mean, median, standard deviation, and 99% confidence interval of the mean for centroid differences.
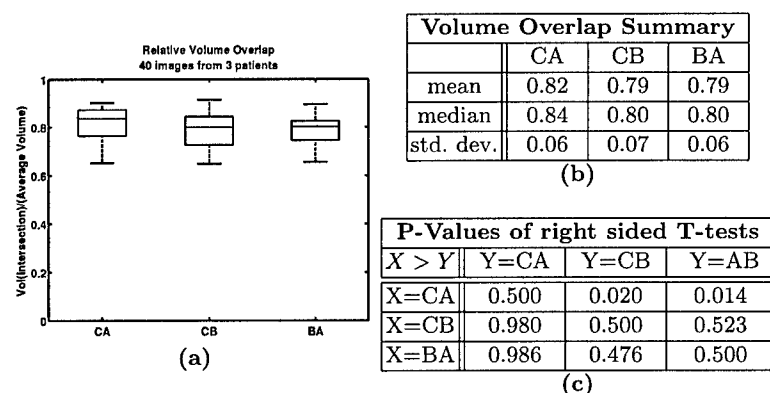
in several ways. For this study, we define the relative volume overlap for two segmentations S_1 and S_2 as

$$\Theta(S_1, S_2) = \frac{\text{Volume}(S_1 \cap S_2)}{\left(\frac{\text{Volume}(S_1 \cup S_2) + \text{Volume}(S_1 \cap S_2)}{2} \right)}. \tag{1}$$

Figure 6 (a) shows a box and whisker plot of the relative volume overlap for the CA, CB, and BA comparisons. To statistically quantify the difference between the relative volume overlaps of the three segmentations A, B, and C, we performed right sided t-tests with the alternative hypothesis $X > Y$. Figure 6, panel (c), reports the P-values of these tests. It can be seen from the table that the volume overlap measures for the CA comparisons are significantly higher than the volume overlap measure for the manaul rater comparison BA. There is also no statistically significant difference between the relative volume overlaps from the CB comparison with the two manual raters. Also note that the automatic segmentations have a significantlly better overlap with rater A than with rater B. This is expected as the planning image was segmented by rater A.

4   Conclusion

We have presented an approach for automatically quantifying organ motion for adaptive radiation therapy of the prostate. This method extends deformable image registration to accommodate bowel gas, which creates image regions where no correspondence exists. We statistically analyzed the accuracy of our automatic method against the standard of manual inter-rater variation. We showed that for centroid and volume overlap of the prostate, the automatic method is statistically indestinguishable from human raters. We are currently working on applying our method to evaluate the clinical effect of organ motion by measuing effective delivered dose and biological effect.

Relative Volume Overlap
40 images from 3 patients

Vol(Intersection)/(Average Volume)

CA CB BA

(a)

Volume Overlap Summary			
	CA	CB	BA
mean	0.82	0.79	0.79
median	0.84	0.80	0.80
std. dev.	0.06	0.07	0.06

(b)

P-Values of right sided T-tests			
$X > Y$	Y=CA	Y=CB	Y=AB
X=CA	0.500	0.020	0.014
X=CB	0.980	0.500	0.523
X=BA	0.986	0.476	0.500

(c)

Fig. 6. (a) Relative volume overlap as measured by Equation 1. The horizontal lines on the box plots represent the lower quartile, median, and upper quartile values. The whiskers show the extent of the rest of the data. (b) Volume overlap summary statistics. (c) P-value results of right sided t-test comparing the relative volume overlaps between the various rators.

References

1. Yan, D., Lockman, D., Brabbins, D., Tyburski, L., Martinez, A.: An off-line strategy for constructing a patient-specific planning target volume in adaptive treatment process for prostate cancer. International Journal of Radiation Oncology*Biology*Physics **48** (2000) 289–302
2. van Herk, M., Bruce, A., Guus Kroes, A.P., Shouman, T., Touw, A., Lebesque, J.V.: Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. International Journal of Radiation Oncology*Biology*Physics **33** (1995) 1311–1320
3. Ketting, C.H., Austin-Seymour, M., Kalet, I., Unger, J., Hummel, S., Jacky, J.: Consistency of three-dimensional planning target volumes across physicians and institutions. International Journal of Radiation Oncology*Biology*Physics **37** (1997) 445–453
4. Joshi, S., Lorenzen, P., Gerig, G., Bullitt, E.: Structural and radiometric asymmetry in brain images. Medical Image Analysis **7** (2003) 155–170
5. Miller, M.I., Joshi, S.C., Christensen, G.E.: Large deformation fluid diffeomorphisms for landmark and image matching. In Toga, A.W., ed.: Brain Warping. Academic Press (1999)
6. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. IEEE Transactions On Image Processing **5** (1996) 1435–1447
7. Christensen, G.E., Joshi, S.C., Miller, M.I.: Volumetric transformation of brain anatomy. IEEE Transactions on Medical Imaging **16** (1997) 864–877
8. Amenta, N., Choi, S., Kolluri, R.K.: The power crust. In: ACM Symposium on Solid Modeling and Applications. (2001) 249–260

APPENDIX B

Large deformation 3D image registration in image-guided radiation therapy

Mark Foskey, Brad Davis, Lav Goyal, Sha Chang, Ed Chaney, Nathalie Strehl, Sandrine Tomei, Julian Rosenman and Sarang Joshi

Department of Radiation Oncology, University of North Carolina

E-mail: mark_foskey@unc.edu

Abstract.   In this paper, we present and validate a framework, based on deformable image registration, for automatic processing of serial 3D CT images used in image-guided radiation therapy. A major assumption in deformable image registration has been that, if two images are being registered, every point of one image corresponds appropriately to some point in the other. For intra-treatment images of the prostate, however, this assumption is violated by the variable presence of bowel gas. The framework presented here explicitly extends previous deformable image registration algorithms to accommodate such regions in the image for which no correspondence exists.

We show how to use our registration technique as a tool for organ segmentation, and present a statistical analysis of this segmentation method, validating it by comparison with multiple human raters. We also show how the deformable registration technique can be used to determine the dosimetric effect of a given plan in the presence of non-rigid tissue motion. In addition to dose accumulation, we describe a method for estimating the biological effects of tissue motion using a linear-quadratic model. This work is described in the context of a prostate treatment protocol, but it is of general applicability.

1. Introduction

In radiation cancer therapy, the problem of organ motion over the course of treatment is becoming more urgent as techniques for conformal therapy improve. These techniques, such as intensity modulated radiation therapy (IMRT), offer important benefits: With high gradients between the region receiving a therapeutic dose and surrounding regions, it is possible, in principle, to increase the prescribed dose to the tumor while reducing the dose to critical organs. The problem with these high gradients is that organ location varies between treatment days, because of both setup error and internal changes such as bowel and bladder filling. With high dose gradients, relatively little organ motion is required to bring parts of the tumor outside of the therapeutic region, or to bring healthy critical tissues in. Both forms of tissue misplacement can harm the patient, in the one case by failure of local control, and in the other, by toxicity to normal tissue. There are now in-the-treatment-room imaging methods, such as cone beam CT and CT-on-rails, that enable image guided radiation therapy as a way to meet this challenge. However, there remains a pressing need for automatic techniques to translate these images into useful information about organ location and likely treatment effectiveness.

The traditional approach to the problem of organ motion has been to specify a margin around the clinical target volume (CTV) to create the planning target volume (PTV). The goal of the margin is to achieve a specified confidence level, interpreted as the probability, at a given treatment session, that actual tumor is contained entirely within the PTV. Work by Goitein and Busse (1975) and Goitein (1985, 1986) suggests that a confidence level of 95% is required. Typically, the size of the margin is expressed as a single parameter, its width, which is based on studies of organ motion across populations of patients. Sometimes the width is reduced near critical structures. For instance, with prostate cancer, the size of the margin may be set to 1 cm, with a reduction to 6 mm toward the rectum (Happersett *et al* 2003).

This simple construction of the PTV relies on two assumptions that have been necessitated by technical limitations in treatment planning and delivery. The first assumption is that organ motion has the same statistical properties for different patients, so that the variance in organ position for a single patient will be equal to that computed previously for a population of patients. The second assumption is that organ motion is statistically the same for all parts of the organ.

To avoid having to make the first assumption, Yan *et al* (1997) introduced the framework of adaptive radiation therapy (ART), in which organ motion for the individual patient is measured over the course of treatment, and the PTV is modified once the amount of motion for that patient has been estimated with sufficient confidence. In their work, the position variation is expressed as a single parameter, a 95% confidence radius for the position of the tumor isocenter, thus still making the assumption that motion is uniform across the relevant organs.

To account for motion that is not uniform, in which organs deform and move relative to one another, a more sophisticated analysis of images is necessary. Recent

computational advances have enabled the emergence of a discipline called computational anatomy (Grenander and Miller 1998) with the principal aim of developing specialized mathematical and software tools for the precise mathematical study of anatomical variability. Within computational anatomy, *deformable image registration* techniques have proved to be effective in the study of anatomical variation (Davatzikos 1996, Christensen *et al* 1997, Csernansky *et al* 1998, Joshi *et al* 1997, Thompson and Toga 2002).

In the framework of computational anatomy, this paper presents a comprehensive approach for automatic processing of 3D CT images acquired during image guided radiation therapy. Deformable image registration is the key to the approach, making it possible to establish a correspondence between points in images taken on different days. Such a correspondence is useful in two key ways: It facilitates automatic organ segmentation, and it makes it possible to calculate the dosimetric effects of nonrigid tissue motion.

The need for careful repeated segmentations has been one of the major limitations for the widespread application of ART and other image guided techniques. Although careful manual segmentation techniques remain the standard of practice, a full manual segmentation of the intra-treatment CT images is time consuming, expensive and not practical in a routine clinical setting. Moreover, manual segmentation introduces uncertainties associated with variability both between and within raters. Two European studies that focused on user-guided tumor segmentation found large inter-user variabilities for well circumscribed lesions (Leunens *et al* 1993, Valley and Mirimanoff 1993).

The dosimetric analysis of tissue motion has the potential to permit more sophisticated ART planning than is currently being pursued (Birkner *et al* 2003). A number of groups have studied the dosimetry of rigid patient motion (Booth and Zavgorodni 2001, Booth 2002, Unkelbach and Oelfke 2004), and there has also been some work in dosimetric analysis of deforming tissue (Schaly *et al* 2004, Yan *et al* 1999). The registration algorithm we describe here differs from previous work in that it provides a fully automated means of performing dose accumulation that can handle large deformations.

In the context of radiotherapy of the prostate or cervix, several deformable image registration methods are currently being investigated for alignment of serial CT data sets. Schaly *et al* (2004) use an approach based on thin-plate splines (Bookstein 1989) for matching CT volumes, where homologous points are chosen from manually drawn organ segmentations. They use the resulting displacement fields to measure cumulative dose over multiple fractions for prostate cancer patients. Christensen *et al* (2001) reported registration of serial CT images for patients undergoing treatment for cervix cancer. Their method matches the boundaries of the bladder, rectum, and vagina/uterus, which are first manually segmented in the planning and treatment images. As with our work, they use a viscous-fluid model that accommodates large deformation changes in the anatomy. Wang *et al* (2005) register CT volumes using a method similar

to the demons algorithm of Thirion (1998). Their method employs a voxel-based driving force motivated by optical flow and a Gaussian regularization kernel. They provide an example of automatic segmentation of a treatment image using the resulting deformation fields. Lu *et al* (2004) present a deformable registration technique based on the minimization of an energy functional that combines an image matching term with a smoothness measure on the resulting deformation field. However, none of these studies address the problem of bowel gas for deformable registration of CT images. Also, while some authors have presented validation studies based on known transformations or phantoms, to our knowledge none have presented a large scale analysis of the accuracy of their methods for automatic segmentation of treatment images based on manual contours.

To give background for what follows, we briefly describe the ART protocol (adapted from Yan *et al* 2000) that we use in our regular prostate care. The fundamental purpose is to use a planning target volume (PTV) that reflects the typical organ motion of the particular patient. Rather than attempting to determine that motion prior to treatment, we use a conventional plan during the first five treatment days, at the same time acquiring a registered CT scan each day. After the fifth treatment day, we construct a new PTV by placing a margin around the approximate convex hull of the CTVs from the first five treatment days, and then generate a new plan, this time using IMRT, based on the new PTV. For the remainder of the treatment period, images are acquired twice weekly to indicate whether further adjustments may be necessary. For each image, the patient is first set up for treatment using crosshair tattoos that are aligned with laser fiducials. Then CT-visible skin markers (2.3-mm "BBs") are placed at the locations marked by the lasers, so that the treated isocenter is indicated on the scan. In a future paper we will assess the effectiveness of this protocol in our practice, using the dosimetric techniques described in this paper.

Shown in figure 1 is a visualization of the organ motion over the course of treatment for 9 patients treated in our clinic using the ART protocol. The internal organ motion of the prostate shown in the images was estimated using manual segmentations of intra-treatment CT images acquired by the CT-on-rails system.

The rest of the paper will be organized as follows. In Section 2 we explain the registration algorithms that we use. In Section 3 we explain how we use deformable registration as a tool for segmentation, and evaluate the reliability of the resulting segmentations. In Section 4 we explain dosimetric applications of our algorithms, and we conclude in Section 5.

2. Deformable image registration

The key to our approach is the measurement of organ motion by means of deformable image registration. We interpret the term "organ motion" broadly, to include setup error and any internal tissue displacement or deformation. We measure organ motion by comparing a CT image taken at planning time to a *treatment image* taken immediately
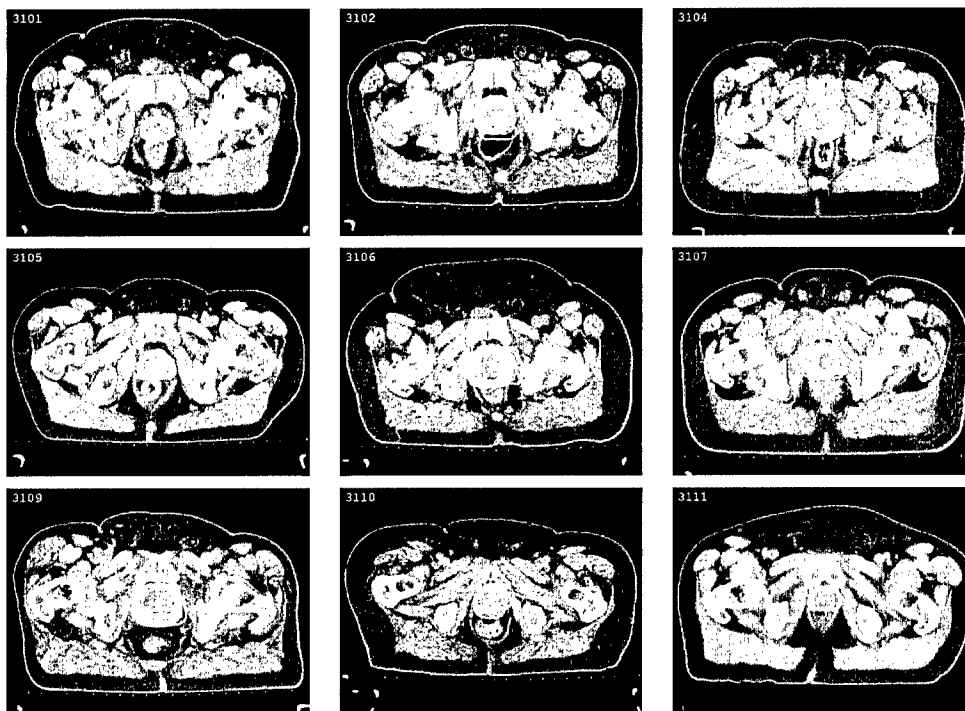
Figure 1. Visualization of prostate motion over the course of treatment for 9 patients involved in our study. White contours, superimposed on an axial slice of each patient's planning image, indicate the actual location of the prostate on each treatment day. These contours are taken from manual segmentations of treatment images. The discrepancies between the contours exhibit the effect of setup error and organ motion on the prostate position. Note that different patients exhibit different amounts of prostate motion; compare the close contour agreement for patient 3101 with the wide contour variability for patient 3109. For some patients (3102, 3109) motion is primarily noticeable in the anterior-posterior direction; for other patients (3106, 3107) motion is primarily noticeable in the lateral direction.

before a given treatment, both of which are acquired using a Siemens Primatom system that provides a CT scanner sharing a table with the treatment machine. If there were no organ motion, the planning image and all the treatment images would be the same, except for noise from the imaging device. However, because there is organ motion, these images will differ, and the difference characterizes the motion (figure 2).

Figure 3 compares a difference image between two unregistered images (aligned as treated) to the difference image for the same two images after registration has been performed. We have understood the motion when we can tell, for each point in the planning image, which point in the treatment image it corresponds to. In this way organ motion and image registration are linked—we can understand organ motion if we can estimate image correspondence. Once image correspondence is established, contours of structures such as the tumor body can be transformed, and other detailed analysis of the changes can be done. The purpose of this section is to explain the registration algorithms we use to establish the correspondence.
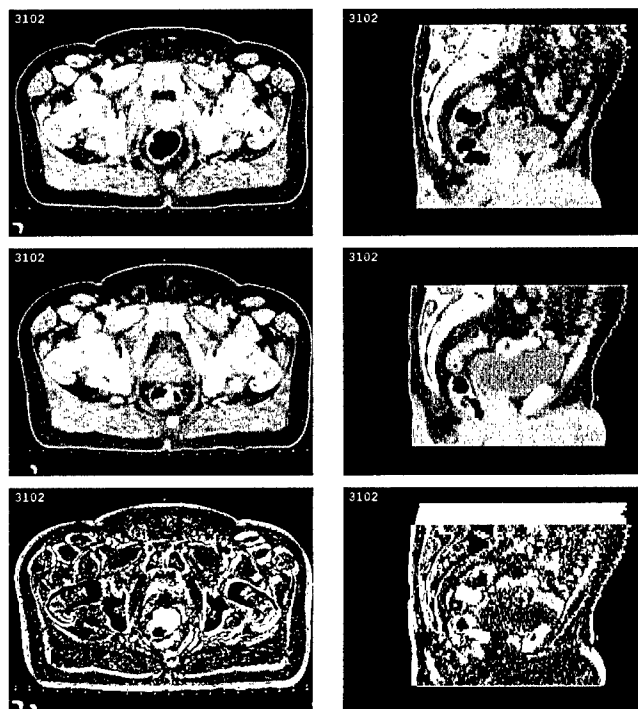
Figure 2. First row: axial and sagittal slices from the planning image of patient 3102. Second row: the same slices (with respect to the planning image coordinate system) taken from a treatment image. Third row: the voxelwise absolute difference between the planning and treatment images. Black represents perfect intensity agreement, which is noticeable in the interior of the bones and outside the patient. Brighter regions, indicating intensity disagreement, are especially apparent: (1) in regions where gas is present in one image and absent in the other, (2) around the bladder which is large on the treatment day compared to the planning day, (3) uniformly along boundaries with high intensity gradient, indicating a global setup error such as a translation.
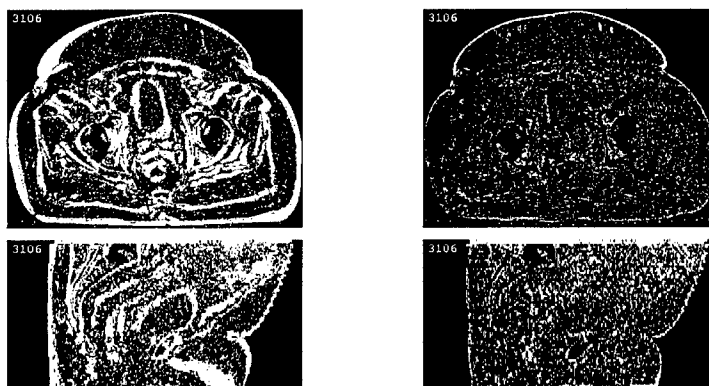
Difference Before Registration   Difference After Registration

Figure 3. Difference images comparing a planning to a daily image before and after deformable registration.

In our discussion we use the term *tissue voxel* to refer to a volume of tissue small enough to be considered as a single point for the purposes of analysis. We view an image as a function $I(x)$ from a domain $V \subset \mathbb{R}^3$ to $\mathbb{R}$, so that $I(x)$ is the intensity of the image at the point $x \in V$. Then the image correspondence can be expressed as a function $h \colon V \to V$, called a *deformation field*. For $x \in V$, $h(x)$ is the point in the treatment image that corresponds to x in the planning image. To the extent that the image registration corresponds to the tissue motion, $h(x)$ is the location, at treatment time, of the tissue voxel originally at x. We find $h(x)$ by approximately minimizing an energy term

$$E(h) = \int_V \left(I_{\mathrm{P}}(x) - I_{\mathrm{T}}(h(x)) \right)^2 \mathrm{d}x, \qquad (1)$$

subject to an appropriate regularity condition. It makes sense to minimize the squared differences of image intensities directly because the CT intensities (expressed in Hounsfield units) have direct physical meaning. The fact that the same machine is used to acquire all images reduces the chance of calibration error.

We decompose the motion into two components, a global rigid transformation (translation and rotation) followed by a deformation that allows the soft tissue to align. This decomposition improves performance since the rigid alignment is fast and accounts for a large portion of the image misalignment. It also makes sense from a clinical perspective since the rigid misalignment corresponds closely to patient setup error and can thus be used to provide guidance for improving setup techniques.

2.1. Rigid Motion

We have used both translation and general rigid motion in our work. For clarity we will explain our algorithm for translation first, after which we will explain the changes needed for general rigid registration. The algorithm can also be modified for general affine registration.

In the case of translation, we want to minimize the energy E subject to the condition that $h(x)$ is of the form $x + \tau$ for some translation vector τ. Thus (1) becomes

$$E(\tau) = \int_\Omega \left(I_{\mathrm{P}}(x) - I_{\mathrm{T}}(x + \tau) \right)^2 \mathrm{d}x$$

Following Joshi *et al* (2003), we use a quasi-Newton algorithm to minimize $E(\tau)$, constructing a sequence $\{\tau_k\}$ such that $E(\tau_k)$ converges to a local minimum. Let $\tau_{k+1} = \tau_k + \Delta\tau$; we will derive a formula for $\Delta\tau$. For convenience, write $x' = x + \tau_k$. If we expand $I_{\mathrm{T}}(x + \tau_{k+1}) = I_{\mathrm{T}}(x' + \Delta\tau)$ in a first order Taylor series about $x' + \Delta\tau$, we get

$$E(\tau_{k+1}) \approx E_{\mathrm{approx}}(\Delta\tau)$$
$$= \int_\Omega \left(I_{\mathrm{P}}(x) - I_{\mathrm{T}}(x') + \nabla I_{\mathrm{T}}(x') \cdot \Delta\tau \right)^2 \mathrm{d}x.$$

At each step in the iteration we minimize $E_{\text{approx}}(\Delta\tau)$ by setting its gradient to 0 and solving for $\Delta\tau$, getting

$$\Delta\tau = -\left(\int_{\Omega} \nabla I_{\text{T}}(x')\nabla I_{\text{T}}(x')^T \, dx\right)^{-1} \int_{\Omega} \left(I_{\text{P}}(x) - I_{\text{T}}(x')\right)\nabla I_{\text{T}}(x') \, dx. \tag{2}$$

In a more general setting, we consider a transformation h that depends on a parameter vector a as well as x, so that we may write $h = h_a(x)$. We then want to find Δa. The expression $I_{\text{T}}(h_a(x))$ is a function of both x and a, and, in the same way that (2) was derived, we find that

$$\Delta a = -\left(\int_{\Omega} \nabla_a I_{\text{T}}(h_a(x))\nabla_a I_{\text{T}}(h_a(x))^T \, dx\right)^{-1}$$
$$\times \int_{\Omega} \left(I_{\text{P}}(x) - I_{\text{T}}(h_a(x))\right)\nabla_a I_{\text{T}}(h_a(x)) \, dx. \tag{3}$$

In the case that h is an affine transformation, $\nabla_a I_{\text{T}}(h_a(x))$ can be expressed conveniently in the following way. We define the parameter vector a by

$$a = [\ A_{11}\ \ A_{12}\ \ A_{13}\ \ A_{21}\ \ \ldots A_{32}\ \ A_{33}\ \ \tau_1\ \ \tau_2\ \ \tau_3\ ]^T.$$

We then define, for any point $x = (x_1, x_2, x_3)$,

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & x_1 & x_2 & x_3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x_1 & x_2 & x_3 & 0 & 0 & 1 \end{bmatrix},$$

so that $Ax + \tau = Xa$. With this convention, $\nabla_a I_{\text{T}}(h_a(x)) = \nabla_x I_{\text{T}}|_{h_a(x)}^T X$, which can be easily computed and used in (3).

For *rigid* registration, as opposed to affine, at each iteration we perform the the same step as for affine registration, and then project the resulting matrix to the space of rotation matrices by taking the orthogonal matrix factor of the polar decomposition.

2.2. Deformation

In the case of large deformation registration, rather than constraining h by requiring that it be expressed in a specific form, we modify the energy functional by adding a regularity term that quantifies how severely h deforms the image. Thus we get

$$E(h) = \int_V \left(I_{\text{P}}(x) - I_{\text{T}}(h(x, t))\right)^2 \, dx + E_{\text{reg}}(h).$$

In Bayesian terms, the first term is a likelihood estimate, and the second is a kind of prior on the space of transformations. The key difficulty in this kind of registration is to find a prior that permits large deformations but not arbitrary rearrangements of voxels. The solution that we adopt was first detailed by Christensen *et al* (1996) and further developed by Joshi and Miller (2000). The idea is to introduce a time parameter t and define a function $h(x, t)$ such that $h(x, 0) = x$ and $h(x, t_{\text{final}}) = h(x)$ is the desired

deformation field that aligns I_P and I_T. We construct h as the integral of a time-varying velocity field:

$$h(x,t) = x + \int_0^t v(h(x,s),s)\,\mathrm{d}s,$$

and we define

$$E_{\text{reg}}(h) = \int_{V,t} \|L_{\text{reg}}v(x,t)\|^2\,\mathrm{d}x\,\mathrm{d}t$$

where L_{reg} is some suitable differential operator. In this way, the size of E_{reg} is not directly based on the difference between $h(x)$ and x, which would tend to prevent large deformations. In the context of landmark-based image registration, Joshi and Miller (2000) show that this method, with proper conditions on L_{reg}, produces a diffeomorphism (i.e. differentiable with a differentiable inverse). As a result, each position x in the planning image corresponds to a unique position in the treatment image, and no tearing of tissue is allowed.

Optimization of the resulting functional $E(h)$ is computationally intensive, since the velocity vector fields for all time steps must be optimized together Miller *et al* (2002), Beg *et al* (2005). Therefore we follow a greedy approach. At each time step, we choose the velocity field that improves the image match most rapidly, subject to the smoothness prior. Precisely, for each t we minimize

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_x \Big(I_P(x) - I_T(h(x,t) + sv(h(x,t),t))\Big)^2 \mathrm{d}x\Big|_{s=0} + \int_x \|L_{\text{reg}}v(x,t)\|^2\,\mathrm{d}x.$$

After evaluating the derivative and solving the resulting variational problem, we find that v must satisfy the differential equation

$$(I_P(x) - I_T(h(x,t)))\,\nabla I_T(h(x,t)) = Lv(x,t), \tag{4}$$

where L is a differential operator proportional to $(L_{\text{reg}})^\dagger L_{\text{reg}}$.

A number of choices of L are reasonable, depending on the desired behavior of the algorithm. We choose $L = \alpha\nabla^2 + \beta\nabla\nabla\cdot + \gamma$, a choice motivated by the Navier-Stokes equations for compressible fluid flow with negligible inertia. If we interpret v as the velocity field of a fluid, then the left hand side of (4) represents an image force exerted on each point in the fluid domain. The right hand side of the equation expresses the resistance to flow. This notional fluid has the nonphysical property that it resists compression (and dilation) inelastically, so that volume can be permanently added or removed in response to image forces. Also, the γ term, which can be thought of as a "body friction" term, ensures that L is a positive definite differential operator, and hence invertible (Joshi and Miller 2000).

To compute $h(x)$, we integrate the resulting velocity field forward in time until the change in image match between successive time steps drops below a threshold. At each time step we find v, using the fast Fourier transform, by explicitly inverting L in the frequency domain. In the context of landmark-based image registration, Joshi and Miller (2000) show that this method, with proper conditions on L, does produce a

diffeomorphism. To make sure that Euler integration, being discrete, does not introduce singularities, we choose a step size such that the largest distance moved by a voxel between successive time steps is less than the inter-voxel spacing.

2.3. Bowel Gas

In images of the pelvic region, one problem that arises in deformable image registration is associated to the presence of bowel gas. Regions of gas appear as black blobs surrounded by gray tissue (see figures 2 and 4). Typically, there will not be gas at the same location in the intestine for different images, and in that case there is no reasonable diffeomorphism between the domains of the two images. That is, if $x \in V$ is in a region containing gas in the planning image, and there is no intestinal gas in the same part of the treatment image, then there is no location in the treatment image that naturally corresponds to x, and thus no reasonable value for $h(x)$. Solid bowel contents do not produce the same difficulty because they do not contrast greatly with the inner wall of the bowel, and are therefore handled by the compressibility of the fluid flow model.


(a)   (b)
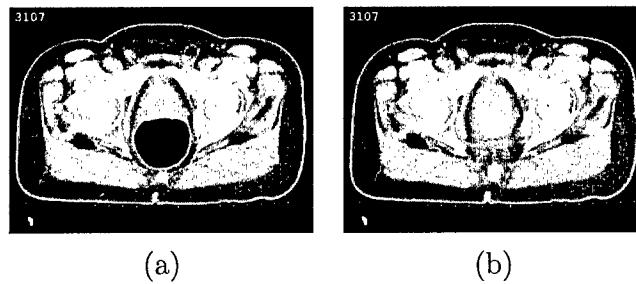
Figure 4. Example of gas deflation. Panel (a) shows an axial slice of a treatment image containing a large region of bowel gas. Panel (b) shows the same image after automatic gas deflation. This deflated image can be accurately registered using deformable image registration.

To resolve the problem of gas, we process each image exhibiting the problem to shrink the gassy region to a point, using a variation of our image deformation algorithm. In doing so, we do not aim to simulate the true motion of the tissue but to deflate the gas so that the image can be accurately registered. The algorithm is defined as follows. We first threshold the image so that gas appears black and tissue appears white, which is possible since the contrast between gas and surrounding tissue is very high in CT images. We then refine this binary segmentation by a morphological opening, contracting and then dilating the gas region, which eliminates small pockets of gas. Using the refined segmentation, we compute a deformation field just as for general deformable registration, by integrating an evolving velocity field $v(x, t)$ to get a deformation field $h(x, t)$. In this case, the velocity field is computed using the equation

$$\nabla I(h(x, t)) = Lv(x, t), \tag{5}$$

with L as in (4). This causes the boundaries of the gas volumes to shrink towards the middle, as if deflating a balloon. Figure 4 shows an axial slice of a treatment image before and after gas deflation.
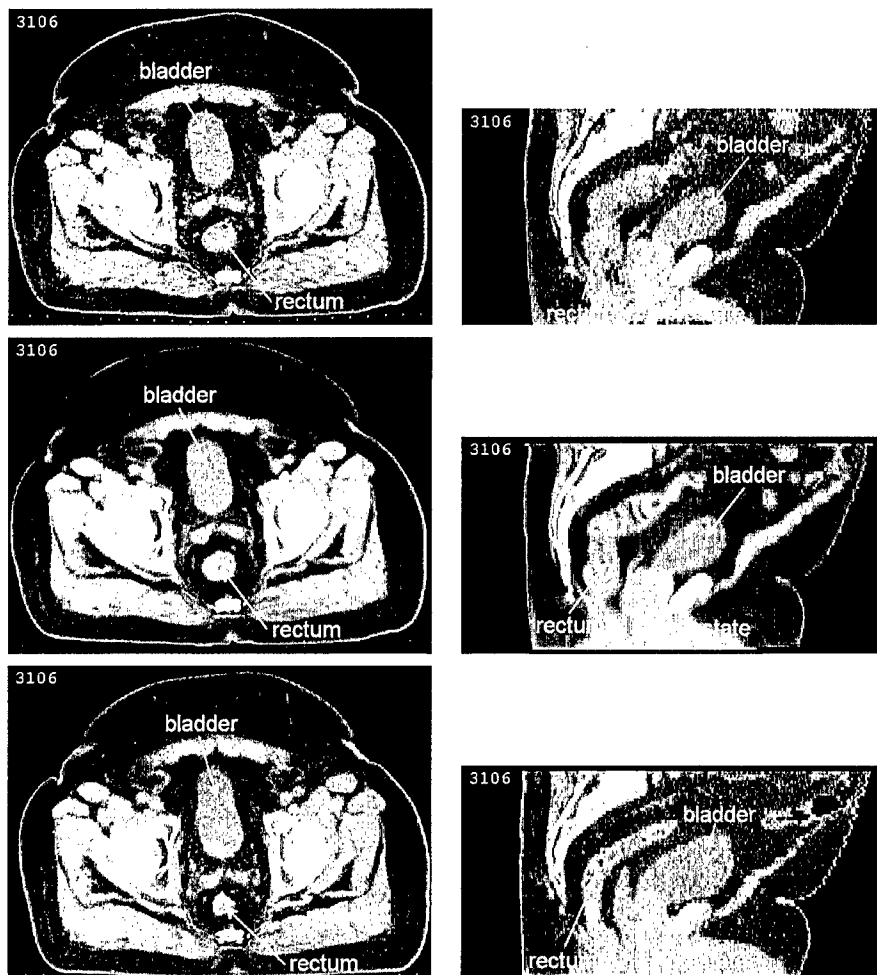
Figure 5. Example of deformable image registration. The first and last rows show axial and sagittal slices of the planning and treatment images. The second row shows the treatment image after deformable image registration, which brings the treatment image into alignment with the planning image. The improvement in soft tissue correspondence suggests that the registration procedure accurately captures internal organ motion. Note how the changes in size and shape of the bladder and rectum are accounted for.

2.4. The Composite Transformation

We now describe how we combine the translation registration, the general fluid registration, and the gas deflation computation to calculate a single transformation from a planning image I_P to a treatment image I_T. We first perform the rigid translation to align the bones as well as possible. For this rigid registration, we choose an intensity

window such that relatively dense bone appears white (maximum intensity), and other tissue appears black. We use a region of interest that includes the medial portion of the pelvis and excludes the femur outside the acetabulum. This computation gives us a translation vector τ.

We then apply the deflation algorithm to I_P and I_T to get two new images $I_{P\text{-}defl}$ and $I_{T\text{-}defl}$, with associated deformation fields $h_{P\text{-}defl}$ and $h_{T\text{-}defl}$ such that $I_{P\text{-}defl}(x) = I_P(h_{P\text{-}defl}(x))$, and similarly for $I_{T\text{-}defl}$. Finally, we apply deformable registration to $I_{P\text{-}defl}$ and $I_{T\text{-}defl}$, yielding the deformation field $h_{TP(defl)}$. Then the full deformation field warping I_T to the space of I_P is given by

$$h_{TP}(x) = h_{T\text{-}defl}(h_{TP(defl)}(h_{P\text{-}defl}^{-1}(x))) + \tau.$$

Accordingly, the point x in the planning image corresponds to the point $h_{TP}(x)$ in the treatment image. This sequence of transformations can be represented as follows:

$$V_P \xrightarrow{h_{P\text{-}defl}^{-1}} V_{P\text{-}defl} \xrightarrow{h_{TP(defl)}} V_{T\text{-}defl} \xrightarrow{h_{T\text{-}defl}} V_{T\text{-}align} \xrightarrow{+\tau} V_T.$$

2.5. Multiscale Registration Implementation

For both rigid and deformable registration we use multiscale techniques to improve efficiency. We resample the images to 1/2 and 1/4 their original resolutions, and then apply our registration algorithm to the coarsest image first, using the result to initialize the algorithm on the next finer image. In the case of deformable registration, we interpolate the deformation field acquired at one resolution to generate the initialization for the immediately finer stage. The parameters we use for α, β, and γ in the definition of L depend on the coarseness of the scale. The values we have used are shown in table 1.

The runtime for the full registration algorithm is proportional to the size of the images being registered, and is dominated by the gas deflation and deformable registration computations, which require two 3D fast Fourier transforms (FFT) per iteration. Each FFT requires on the order of $n \log n$ floating point operations, where n is the number of voxels in each image. For our experiments, n ranges from 1,164,942 ($81 \times 102 \times 141$) to 7,912,905 ($187 \times 217 \times 195$), depending on the patient.

For deformable registration specifically, the time per iteration, averaged over all patients in our study, is 0.2 sec, 2.0 sec, and 22.7 sec for 1/4, 1/2, and full resolution computations, respectively. Timings for gas deflation are slightly less. These results were obtained on a PC with 4GB of main memory and dual 3GHz Intel Xeon processors (although only one processor was used in the computations).
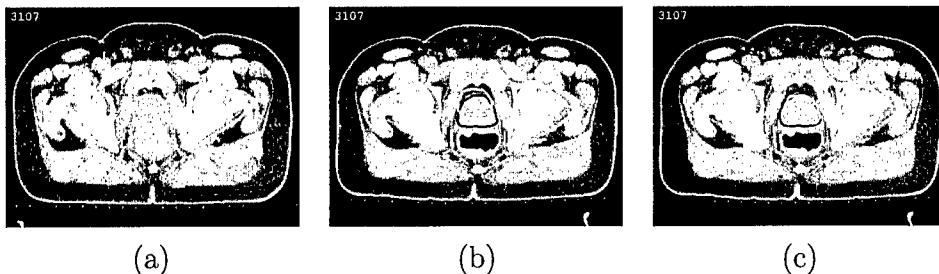
3. Automatic segmentation

The goal of image guidance in radiation therapy is to measure the changes over time of tumor and organs in both location and shape, so that the treatment can be adjusted accordingly. In our current ART practice we use manual contouring of organs for this

Table 1. Parameters used in the regularizing operator $L = \alpha\nabla^2 + \beta\nabla\nabla \cdot + \gamma$.

Scale	α	β	γ	Iterations
Coarse	0.01	0.01	0.001	150
Medium	0.01	0.01	0.001	75
Fine	0.02	0.02	0.0001	25

purpose, but this is problematic because it is time consuming, and because there is considerable variation even when the same individual contours an image repeatedly on different days (Collier *et al* 2003). Instead, using image deformation, it is possible to carry the contours from the planning image to a daily image, deforming them to match the new image. This provides an automatic segmentation of the new image, based on the manual segmentation of the planning image. In this section, we explain our method, and then present a statistical analysis of its accuracy and reliability.

The idea is to use the deformation fields to move the vertices of the contours from their locations in the planning image to the corresponding points in the treatment image (figure 6). This process does not result in a set of planar contours, since vertices will

(a)   (b)   (c)

Figure 6. Example of automatic segmentation using deformable image registration. Panel (a) shows an axial slice of a planning image with the prostate labeled by a white contour. Panel (b) shows the same axial slice (in terms of planning coordinates) from a treatment image. Patient setup error as well as internal organ motion and changes in rectal filling are responsible for the misalignment of the planned prostate position (white) relative to the actual prostate position at treatment time (black, manual segmentation). Panel (c): The same treatment image and manual (black) contour are shown. Deformable image registration is used to estimate the correspondence between the planning and treatment images. The white contour is automatically generated by *deforming* the planning prostate segmentation based on this correspondence. The close agreement of the automatically generated segmentation with the actual prostate position indicate that the deformable image registration accurately captures the prostate motion between planning time and treatment time.

typically be moved out of plane to varying degrees. Therefore, instead of working with the contours directly, we first convert the sequence of contours to a surface model made up of triangles (figure 7) using an algorithm due to Amenta *et al* (2001). Then, we replace each vertex x in the model with $h(x)$, after which we slice the model with planes parallel to the xy axis to generate a new set of contours.

Figure 8 permits a visual assessment of the accuracy of our method. This figure is
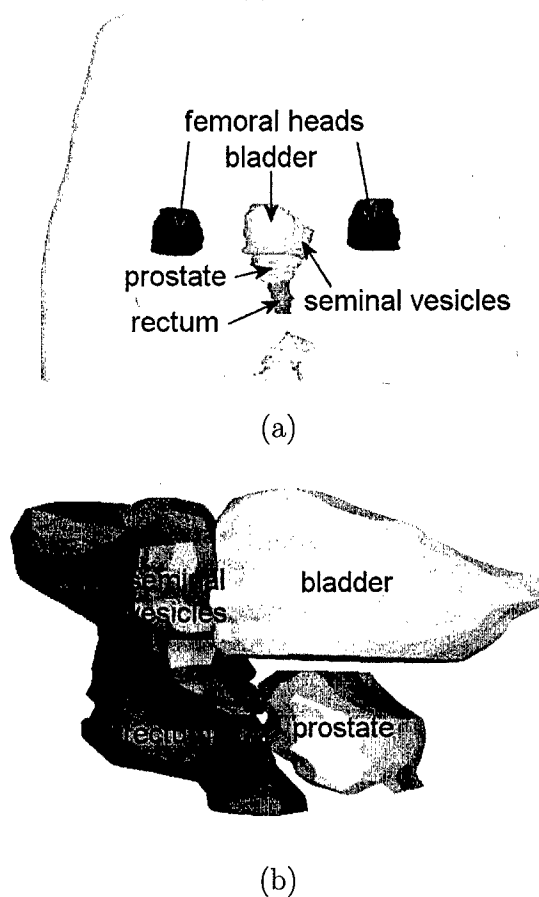
Figure 7. Visualization of organ segmentations. Panel (a) is an anterior view of a 3D rendering displaying segmentations of the skin, prostate, rectum, bladder, seminal vesicles, and femoral heads. Panel (b) shows a lateral view of the prostate, rectum, and bladder of the same patient. The surfaces are constructed by tiling manually drawn contours.

similar to figure 1 except that instead of denoting the actual daily prostate positions, the contours represent the daily prostate positions *deformed* into the space of the planning image. The close agreement of the deformed contours indicates that the image registration algorithm accurately estimates the correspondence between the planning and treatment images along the prostate boundary. Discrepancies between the deformed segmentations are a consequence of both image registration errors and intra-rater variability of the manual, treatment-day segmentations.

Our statistical analysis is based on comparing automatically generated segmentations to manual, hand-drawn segmentations. However, there is appreciable variation in manual segmentation, making it unreasonable to choose a particular manual segmentation as definitive. Groups have reported segmentation variation in a number of contexts, including brain tumors (Leunens *et al* 1993), lung cancer (Valley and Mirimanoff 1993,

Ketting *et al* 1997), and prostate MR (Zou *et al* 2004). Rasch *et al* (1999) reported inter-user variabilities in the segmentation of the prostate in CT and MRI, finding overall observer variation of 3.5mm (1 standard deviation) at the apex of the prostate and an overall volume variation of up to 5% in CT.

Given this inter-rater variability, we assess our method by comparing our automatically generated segmentations with the segmentations from one manual rater (A), and then make the same comparisons between segmentations from a second manual rater (B) and rater A. We judge the accuracy and reliability of the automatic segmentations based on the standard of the measured inter-rater variability.

We have acquired CT scans for a total of 138 treatment days from 9 patients enrolled in our protocol. All of these images have been manually segmented by at least one expert. However,. due to the time-consuming nature of manual segmentation, images from only 4 of these patients have been manually segmented by a second expert. We use images from these 4 patients for the analysis in this section. Eventually we plan to perform the same analysis for all of the patients enrolled in our protocol, and volume overlap statistics for the available segmented organs for all 9 patients are presented in Section 3.2.

The experimental setup is as follows. This study is based on a total of 48 CT images representing 48 treatment days for 4 patients. Each CT scan was collected prior to treatment on the Siemens Primatom scanner mentioned above, with a resolution of $0.098 \times 0.098 \times 0.3$ cm. Each planning image is manually segmented by rater A. Each treatment image is manually segmented twice, once by rater A and once by rater B. For each patient, our method is used to compute the transformations h_i that deformably align the planning image with the treatment image for each day of treatment i. An automatic segmentation is generated for each treatment image by applying h_i to the segmentation in the planning image. We consider our automatic method for producing segmentations as rater C (for "computer").

For each patient and for each treatment day, we make two comparisons: **CA**, automatic segmentation versus segmentation by rater A, and **BA**, comparing manual segmentations by raters B against those by rater A. It should be emphasized that the automatic segmentations are produced by transforming manual planning segmentations produced by rater A, not rater B.

In the rest of this section, we present the results of this experiment when measuring centroid differences and volume overlap of segmentations. We also show radial distance maps which help us understand which regions of the prostate have the largest segmentation differences.

3.1. Centroid Analysis

The centroid of the prostate is especially important for radiation treatment planning and therapy because it is the origin, or isocenter, for the treatment plan. To measure the accuracy of our automatic segmentations with respect to centroid measurement,
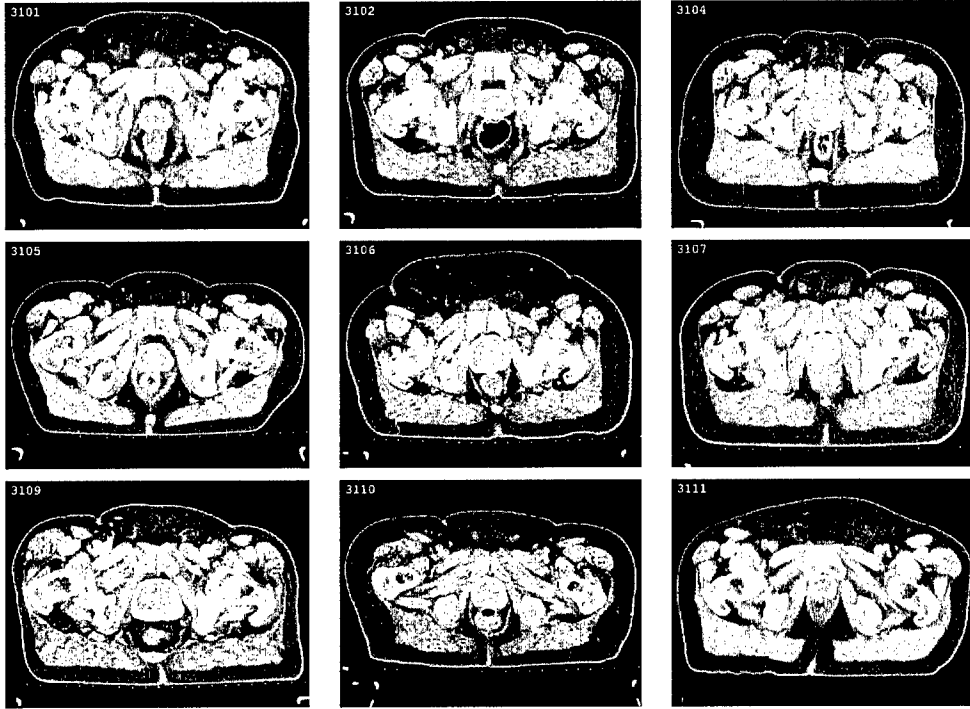
Figure 8. Visualization of the result of image registration algorithm. The images show manual segmentations of each daily image deformed into the space of the planning image using the deformable image registration. The close agreement of the deformed segmentations with the position of the prostate in the planning images provides evidence that the image registration algorithm accurately estimates the correspondence between planning and treatment images along the prostate boundary.

we compare the centroid of each automatic segmentation with the centroid of the corresponding manual segmentation.

First we consider the question: Are the centroids of the automatic segmentations systematically shifted with respect to rater A's segmentation? Let S_A^i, S_B^i, and S_C^i denote the prostate segmentations from raters A, B, and C, respectively, for image i. Let $\mathcal{C}(\cdot)$ be a function that returns the centroid (in $\mathbb{R}^3$) of a segmentation. In order to determine whether the centroids of the automatic segmentations are systematically shifted in any particular direction, we examine the distribution of the centroid differences $\mathcal{C}(S_C^i) - \mathcal{C}(S_A^i)$, $i \in 1, 2, \ldots N$. Likewise, to test for shifts between manual raters A and B, we examine the distribution $\mathcal{C}(S_B^i) - \mathcal{C}(S_A^i)$. Figure 9 shows box-and-whisker plots of these differences for the CA and BA comparisons. The differences in the lateral (X), anterior-posterior (Y), and superior-inferior (Z) directions are measured separately. Summary statistics are provided in Table 2.

It can be seen from this data that there is no significant shift between centroids of the computer generated segmentations and rater A's manual segmentations in the lateral and A-P directions. There *is* a significant shift ($p < 0.001$ for two-tailed, paired t-test) in the sup-inf direction of approximately 0.1 cm, which is one third of the sup-inf
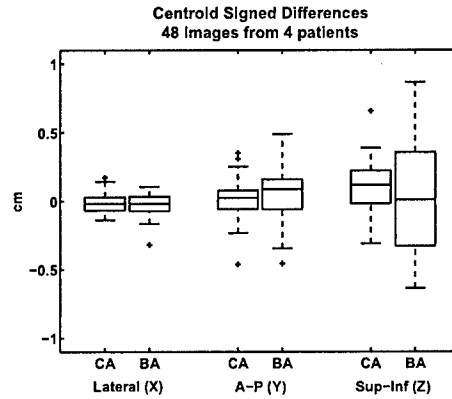
Figure 9. Centroid differences measured in the lateral (X), anterior-posterior (Y), and superior-inferior (Z) directions. The horizontal lines on the box plots represent the lower quartile, median, and upper quartile values. The whiskers indicate the extent of the rest of the data, except that outliers, which fall more than 1.5 times the interquartile range past the ends of the box, are denoted with the '+' symbol.
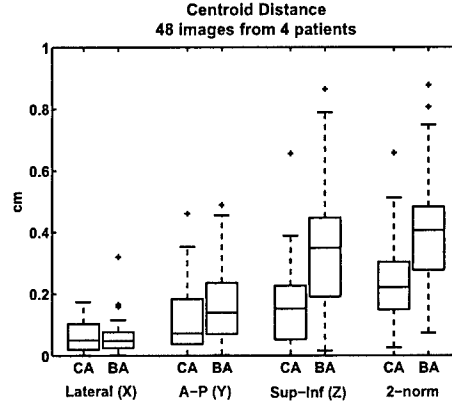
Table 2. Summary statistics for centroid difference distributions $\mathcal{C}(S_C^i) - \mathcal{C}(S_A^i)$ and $\mathcal{C}(S_B^i) - \mathcal{C}(S_A^i)$. The mean, standard deviation, and 95% confidence interval for the mean are reported.

	Lateral (X)		A-P (Y)		Sup-Inf (Z)	
	CA	BA	CA	BA	CA	BA
Mean	-0.01	-0.02	0.01	0.06	0.10	0.02
STD	0.08	0.08	0.15	0.20	0.18	0.38
95% CI	-0.04	-0.05	-0.03	0.00	0.05	-0.09
	0.01	0.00	0.06	0.12	0.15	0.14

image resolution (0.3 cm). In all three directions, the standard deviation of the BA comparisons is as large or larger than the standard deviation of the CA comparisons. Most notably, the standard deviation of the manual rater comparison is twice as large as the standard deviation of the automatic-manual comparison in the sup-inf direction.

Next we examine the magnitude of the centroid differences measured by $\|\mathcal{C}(S_A^i) - \mathcal{C}(S_B^i)\|_2$ and $\|\mathcal{C}(S_A^i) - \mathcal{C}(S_C^i)\|_2$, where $\| \cdot \|_2$ denotes the 2-norm (Euclidean distance). Figure 10 shows box-and-whisker plots of these distances, as well as the component (lateral, A-P, and sup-inf) distances. Summary statistics for these data are presented in table 3. As the distributions of these distances are not approximately normal, we report medians and interquartile ranges as well as means and standard deviations.

It is clear from figure 10 that for both raters C and B the least amount of distance is measured in the lateral (X) direction while the largest distances are measured in the sup-inf (Z) direction. In fact, the sup-inf distance represents a very large part of the

Figure 10. Distance between segmentation centroids measured separately in X, Y, and Z, as well as Euclidean distance.

Table 3. Summary statistics for centroid distance distributions $|\mathcal{C}(S_C^i) - \mathcal{C}(S_A^i)|$ and $|\mathcal{C}(S_B^i) - \mathcal{C}(S_A^i)|$.

	Lateral (X)		A-P (Y)		Sup-Inf (Z)		Euclidean Distance	
	CA	BA	CA	BA	CA	BA	CA	BA
Mean	0.063	0.060	0.111	0.166	0.162	0.331	0.234	0.404
Median	0.050	0.048	0.073	0.141	0.152	0.348	0.221	0.406
Max	0.174	0.320	0.461	0.489	0.656	0.865	0.657	0.877
STD	0.048	0.054	0.100	0.120	0.128	0.193	0.128	0.178
IQR	0.082	0.051	0.145	0.165	0.174	0.256	0.153	0.206

total distance. Recall that the sup-inf resolution is 0.3 cm, as opposed to the finer 0.098 cm resolution in the lateral and A-P directions.

We see that the centroids of the automatically generated segmentations are consistently closer to the centroids of rater A's segmentations than are the centroids for rater B. In the lateral direction, the CA and BA distances are comparable and within image resolution. In the A-P and sup-inf directions, as well as for the Euclidean norm, the median distance for the CA comparisons is approximately half that of the BA comparisons. Furthermore, the median CA distance is within image resolution for these directions while the median BA distance is not. We tested the CA and BA distances for equality of their means using two-tailed, paired t-tests, finding p-values of 0.775 and 0.022 in the lateral and A-P directions and less than 0.0001 in the sup-inf direction and Euclidean norm. The median distance for CA is 0.17 cm less than the median distance for BA.

We conclude that the automatic segmentation method is as accurate for estimating centroids as human raters and, as seen by the error bars and standard deviations, at least as reliable.
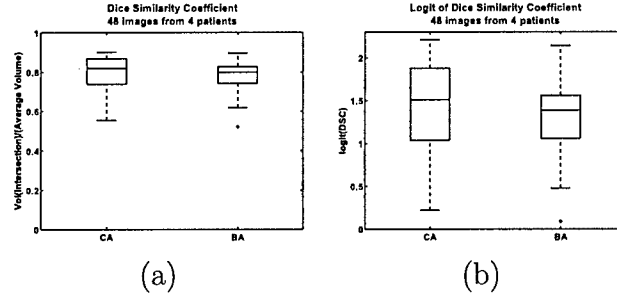
Figure 11. Dice Similarity Coefficient (DSC) and logit DSC for CA and BA comparisons.

3.2. Volume Overlap Analysis

To measure the coincidence between volumetric segmentations of the prostate we use the Dice Similarity Coefficient (DSC) of Dice (1945). For two segmentations, S_1 and S_2, the DSC is defined as the ratio of the volume of their intersection to their average volume:

$$\mathrm{DSC}(S_1, S_2) = \frac{\mathrm{Volume}(S_1 \cap S_2)}{\frac{1}{2}\left(\mathrm{Volume}(S_1) + \mathrm{Volume}(S_2)\right)} \tag{6}$$

The DSC has a value of 1 for perfect agreement and 0 when there is no overlap. A DSC value of 0.7 or greater is generally considered to indicate a high level of coincidence between segmentations (Zijdenbos *et al* 1994, Zou *et al* 2004). The DSC can be derived from the kappa statistic for measuring chance-corrected agreement between independent raters (Zijdenbos *et al* 1994).

Figure 11 (a) shows a box-and-whisker plot of the Dice similarity coefficient for the CA and BA comparisons. The mean DSC for the CA comparisons was 0.80 (STD=0.08) indicating that the automatic segmentations have generally good coincidence with the manual segmentations. The mean DSC for the two manual raters was similar (mean=0.78, STD=0.07).

A similar study, carried out by Zou *et al* (2004), assessed the reliability of manual prostate segmentations in interoprative MR images. They report a mean DSC for manual raters of 0.838. Note that because prostate boundaries are more evident in MR images than in CT images, manual raters are likely to segment MR images more reliably than CT images.

To evaluate the DSC distributions we use the logit of the DSC (LDSC), defined by

$$\mathrm{LDSC}(S1, S2) = \ln\left(\frac{\mathrm{DSC}(S_1, S_2)}{1 - \mathrm{DSC}(S_1, S_2)}\right).$$

Agresti (1990) has shown that for large sample sizes (in the case of our prostate segmentations, the number of voxels is approximately 20 000), LDSC has a Gaussian distribution. Figure 11 (b) shows a box-and-whisker plot of the LDSC values for the CA and BA comparisons. Summary statistics are reported in table 4.

Table 4. Summary statistics for the DSC and LDSC measures.

	DSC(S_1, S_2)		LDSC(S_1, S_2)	
	CA	BA	CA	BA
Mean	0.80	0.78	1.44	1.30
Median	0.82	0.80	1.51	1.39
STD	0.08	0.07	0.49	0.41
IQR	0.13	0.08	0.84	0.50

Table 5. Comparison of automatic segmentation to manual segmenter A via the DSC and LDSC. This is the full set of segmenter-A segmentations that we have processed.

|	DSC(S_1, S_2)		LDSC(S_1, S_2)	
	Prostate	Bladder	Prostate	Bladder
n	76	20	76	20
Mean	0.801	0.816	1.466	1.576
Median	0.825	0.826	1.554	1.557
STD	0.081	0.078	0.494	0.539
IQR	0.121	0.133	0.804	0.034

In order to test for a significant difference between the CA and BA comparisons we performed a paired t-test on the LDSC values. A one-tailed test shows that the DSCs for the CA comparisons are significantly ($p = 0.005$) greater than the DSCs for BA comparisons. Therefore, the automatic segmentations coincide with rater A's segmentations better than a second manual rater.

Table 5 summarizes the manual-versus-automatic comparison for segmenter A only, for all patients that have been processed. After the first five treatment days, the bladder typically was not segmented.

3.3. Radial Distance Maps

Manual segmenters tend to find some portions of the prostate more difficult to segment than others. For instance, in CT there is often little or no apparent contrast between the prostate and bladder. Thus it makes sense to examine segmentation variability as a function of position on the prostate. For two segmentations X and Y of the same image, we can visualize the deviation by choosing the centroid of X as a reference point, and considering, for each ray emanating from the centroid, the distance between the intersection points of the ray with X and Y. For each surface, we choose the first point that the given ray intersects that surface; typically there is only one. This procedure produces a distance for each radial direction, which can be plotted on the surface of a sphere, producing a *radial distance map*. This radial distance map is inspired by that of Mageras *et al* (2004), but we use a slightly different definition. To display the spherical map, we use the cartographic equal-area Mollweide projection. Since the patients are

all scanned in a consistent orientation, different radial distance maps can be compared directly, and average maps can be computed point by point. Figure 12 shows the mean radial distance at each point for the cases analyzed in this section. Notice that the largest variation is generally found in the superior direction, which is consistent with the observed difficulty of detecting the boundary between prostate and bladder.
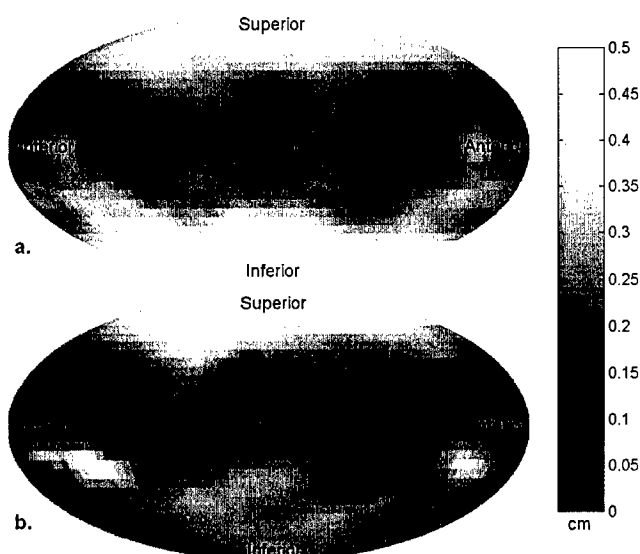

Figure 12. Radial distance maps, for the prostate. Map a: Mean radial distance between segmentations A and B (human raters). Map b: Mean distance between A and C segmentations (human and computer).

4. Dosimetric evaluation of image-guided radiotherapy

The day-to-day effects of organ motion and setup error can be illustrated by computing a DVH based on the observed organ location on each day. Figure 13 shows DVHs for the first four days of treatment for patient 3102 of our protocol. The DVHs were computed by calculating the dose distribution based on the image for the given day, and applying that distribution to the organ segmentation computed by deforming the planning segmentations. Day 2 has a particularly severe cold spot, a fact confirmed by a comparison of the planning and treatment images. In figure 14, contours for the prostate from both the planning day and treatment day 2 are shown, along with isodose lines for 95%, 98%, and 100% of prescribed dose. The top panel shows a full axial slice of the treatment image from Day 2, with an overlaid skin contour from the planning image as an indication of setup accuracy. The bottom two panels show closer views of the prostate from the planning and Day 2 images. In the slice shown, roughly half of the prostate appears to lie outside the 95% dose line.

It is not possible to directly combine a series of DVHs to produce an accurate DVH for the total dose delivered, because each DVH only indicates how great a volume from
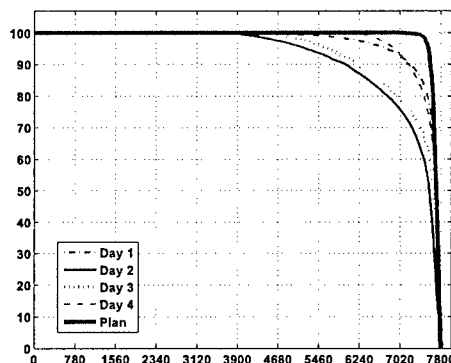
Figure 13. Daily treatment prostate DVHs for each of the first four days, compared to the planned DVH.
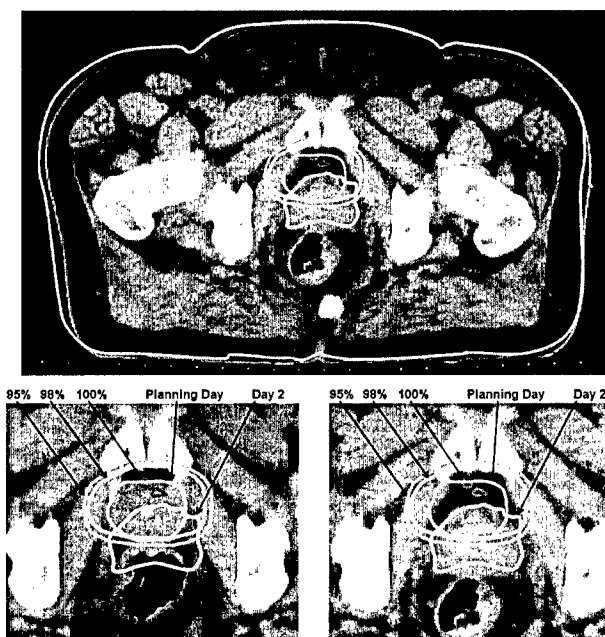
Figure 14. The position of the prostate in patient 3102 at day 2 compared to the time of planning. **Top:** The treatment image from day 2, shown with a skin contour from the planning day. **Bottom left:** Planning image. **Bottom right:** Day 2 image. On all three images, the location of the prostate is shown for both days, along with isodose curves at the 95%, 98%, and 100% level.

a given day received a specific dose. To combine information from different days, one needs to know the daily dose received by each voxel. Bortfeld *et al* (2004) provide a survey on the statistical effects of organ motion on dose distributions, using a rigid model. In the rest of this section we will describe how to assess total delivered dose in actual cases, considering deformation, by applying the displacement fields h computed from deformable image registration. Yan *et al* (1999), Birkner *et al* (2003) and Schaly *et al* (2004) have all described similar approaches, considering both raw and effective

dose. However, their image registration algorithms require either that fiducial points be manually selected in the images, or that all of the images be segmented manually. In addition, none of these methods permit the range of deformations allowed by the fluid model.

4.1. Total Delivered Dose

Let the dose per fraction, as a function of position $x \in V$, be given by $D(x)$. Then the dose received at treatment i, by the tissue originally at x, is given by $D(h_i(x))$, and the total dose received by that tissue voxel over the course of treatment is given by

$$D_{\mathrm{Tot}}(x) = \sum_i D(h_i(x)).$$

Using this formula, we can compute a distribution for total delivered dose, in the frame of reference of the planning day. Using the organ segmentations from the planning image, we can calculate DVHs that correctly reflect the variation in dose distribution over time. Figure 15 shows a series of delivered DVHs for increasing sets of treatment days, all normalized to the same prescription dose of 78 Gy. As expected, the quality of
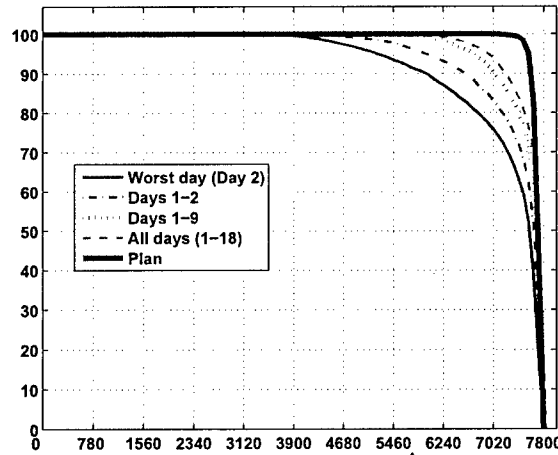
Figure 15. Dose volume histograms for delivered dose, estimated over increasing sets of days. These are compared against the planned dose. All doses are normalized to a prescription dose of 78 Gy.

the DVH improves as the number of treatments being accumulated is increased, and we would expect further improvement given images from all 39 treatment days. But note that the DVH is still quite poor even based on 18 treatments, and that it only improved modestly over the 9-treatment DVH.

4.2. Effective Cumulative Dose

The difficulty with the measure D_{Tot} is that the biological effect does not depend simply on the total dose received, but also on the way it is distributed into fractions. Consider

a volume of cells irradiated to a dose D over a time that is short relative to that required for cell repair to occur. Then the linear quadratic (LQ) model (Fowler 1989) gives the following estimate of the survival fraction SF of the cells in the volume:

$$\mathrm{SF}(D) = \mathrm{e}^{-\alpha D - \beta D^2}$$

Now let $T = (D_1, D_2, \ldots, D_N)$ be a series of *varying* doses separated by time for cell recovery. In our situation, the relevant volume of tissue is a voxel x and, for each i, $D_i = D(h_i(x))$. Assuming that cell proliferation is negligible, the survival fraction for the treatment T will be given by

$$\mathrm{SF}(T) = \prod_i \mathrm{e}^{-\alpha D_i - \beta D_i^2} = \exp\left(\sum_i -\alpha D_i - \beta D_i^2\right).$$

Just as with uniform fractionation, one can construct the Biological Effective Dose (BED; Fowler 1989, Barendsen 1982). The BED is the dose that, if delivered in a series of fractions so small that the β term may be ignored, would kill the same number of cells as the actual dose in question. That is, we define $\mathrm{SF}(\mathrm{BED}) = \mathrm{e}^{-\alpha \cdot \mathrm{BED}}$, and compute the BED for a particular treatment regimen T by setting $\mathrm{SF}(\mathrm{BED}) = \mathrm{SF}(T)$ and solving to obtain

$$\mathrm{BED}(T) = \sum_i D_i + \frac{D_i^2}{\alpha/\beta}.$$

Then, following the analysis for the total delivered dose, we can define the total BED for a tissue voxel x as follows (see also Yan *et al* 1999, Birkner *et al* 2003, Schaly *et al* 2004):

$$\mathrm{BED}_{\mathrm{Tot}}(x) = \sum_i D(h_i(x)) + \frac{D(h_i(x))^2}{\alpha/\beta}. \tag{7}$$

To illustrate, figure 16 compares the delivered BED to that planned, as estimated based on the 18 treatment images. That is, the delivered BED was computed by applying Equation 7 to the appropriate 18 dose distributions and deformation fields, with each distribution based on a prescription dose of 2 Gy/f. The resulting distribution was then normalized to a 78 Gy prescription dose by applying a scale factor of 78/36. As with raw dose accumulation, this estimate does not account for the improvement in the distribution that would result from averaging together a greater number of random motions. For the purposes of illustration we assumed a low α/β value of 1.5 Gy, but this is not far removed from current estimates (Fowler *et al* 2001).

Since a 39-fraction regimen is designed to minimize the effect of the quadratic component on low-α/β tissue, difference between accumulated raw dose and accumulated BED is not great (see Fig. 17). Because of evidence indicating that prostate tumors may have α/β values comparable to healthy tissue, there is now considerable discussion of hypofractionation for prostate cancer (Kupelian *et al* 2002, Brenner 2003, Craig *et al* 2003). Figure 18 shows DVHs of accumulated BED for four values of α/β, assuming a 9-fraction regimen.
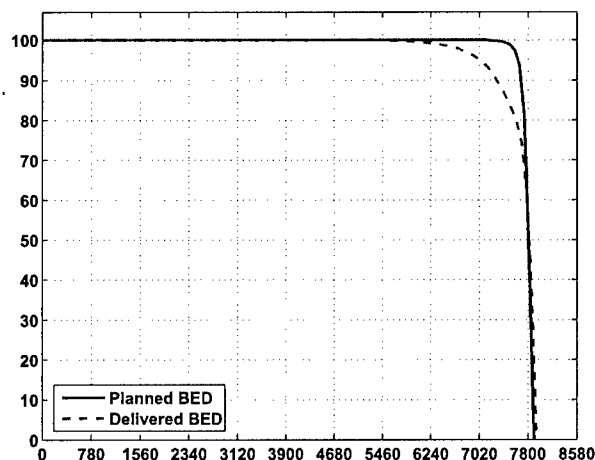
Figure 16. Delivered BED compared to planned BED. Delivered BED is modeled from a sample of 18 out of 39 treatment days. $\alpha/\beta = 1.5$.
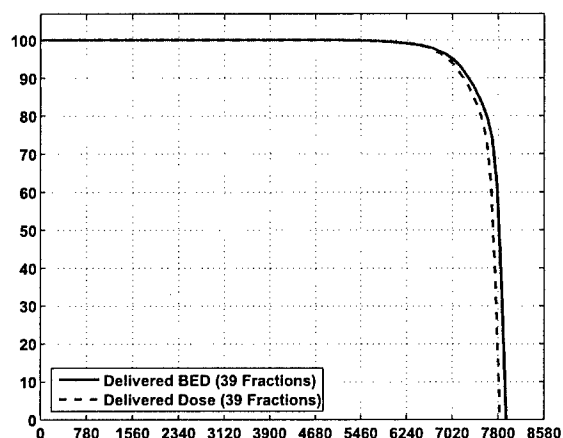

Figure 17. Comparison of nominal and biologically effective dose, as delivered, estimated based on 18 treatment days.

5. Discussion

We have described how large deformation image registration can be used for automatic segmentation and dose accumulation in the course of image guided radiation therapy. Our image registration technique is fully automatic, permits large deformations, and ensures a smooth one-to-one correspondence between two images. We use a variation of the registration method to eliminate bowel gas when it occurs, so that images can be brought into a meaningful correspondence.

For segmentation purposes, we compute the deformation that transforms the planning image to match the daily treatment image, and apply that deformation to the initial manual contours. We have validated this method by comparing the automatic segmentations to manual segmentations produced by the same segmenter who generated
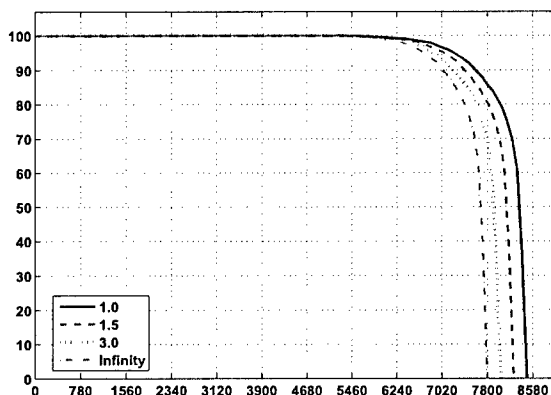
Figure 18. Biologically effective DVHs, assuming four possible values of α/β, and 78 Gy delivered in 9 equal fractions. An actual hypofractionation regime would typically use a smaller total dose, but 78 Gy was used here for comparison with current practice.

the original planning segmentations. Based on centroid difference and the DSC measure of volume overlap, we find that the automatic deformations of a planning segmenter correspond at least as closely to the daily segmentations of the same segmenter, as do daily segmentations by a different individual.

We also show how to use our registration method to estimate the amount of dose delivered to the patient over time, as a function of position within the imaged area. In a single case study we compare daily DVHs to both the planned DVH and to cumulative DVHs, observing that, as expected, the accumulation of multiple fractions tends to improve the correspondence between delivered and planned DVH, though we still find a pronounced difference based on 19 images. We also consider the accumulation of biologically effective dose. For 39 fractions, accumulated BED is very close to accumulated dose, but hypofractionation schemes lead to a greater difference. In the future we intend to apply these dose accumulation measures to assess the effectiveness of protocols both planned and currently in use in our clinic.

Acknowledgments

References

Agresti A *Categorical Data Analysis* Wiley Interscience 1990

Amenta N, Choi S and Kolluri R K The power crust In *ACM Symposium on Solid Modeling and Applications* pages 249–260 2001

Barendsen G W 1982 Dose fractionation, dose rate and iso-effect relationships for

normal tissue responses *International Journal of Radiation Oncology*Biology*Physics* **8** 1981–1997

Beg M F, Miller M I, Trouvè A and Younes L 2005 Computing large deformation metric mappings via geodesic flows of diffeomorphisms *International Journal of Computer Vision* **61** 139–157

Birkner M, Yan D, Alber M, Liang J and Nüsslin F 2003 Adapting inverse planning to patient and organ geometrical variation: Algorithmn and implementation *Medical Physics* **30** 2822–2831

Bookstein F L 1989 Principal warps: Thin-plate splines and the decomposition of deformations *PAMI* **11** 567–585

Booth J T *Modelling the Impact of Treatment Uncertainties in Radiotherapy* PhD thesis Adelaide University 2002

Booth J T and Zavgorodni S F 2001 The effects of radiotherapy treatment uncertainties on the delivered dose distribution and tumour control probability *Australasian Physical & Engineering Sciences in Medicine* **24** 71–78

Bortfeld T, Jiang S B and Rietzel E 2004 Effects of motion on the total dose distribution *Seminars in Radiation Oncology* **14** 41–51

Brenner D J 2003 Hypofractionation for prostate cancer radiotherapy–what are the issues? *International Journal of Radiation Oncology*Biology*Physics* **57** 912–914

Christensen G E, Carlson B, Chao K S C, Yin P, Grigsby P W, N K, Dempsey J F, Lerma F A, Bae K T, Vannier M W and Williamson J F 2001 Image-based dose planning of intracavitary brachytherapy: registration of serial-imaging studies using deformable anatomic templates *International Journal of Radiation Oncology*Biology*Physics* **51** 227–243

Christensen G E, Joshi S C and Miller M I 1997 Volumetric transformation of brain anatomy *IEEE Transactions on Medical Imaging* **16** 864–877

Christensen G E, Rabbitt R D and Miller M I 1996 Deformable templates using large deformation kinematics *IEEE Transactions On Image Processing* **5** 1435–1447

Collier D C, Burnett S S C, Amin M, Bilton S, Brooks C, Ryan A, Roniger D, Tran D and Starkschall G 2003 Assessment of consistency in contouring of normal-tissue anatomic structures *Journal of Applied Clinical Medical Physics* **4** 17–24

Craig T, Moiseenko V, Battista J and Van Dyk J 2003 The impact of geometric uncertainty on hypofractionated external beam radiation therapy of prostate cancer *International Journal of Radiation Oncology*Biology*Physics* **57** 833–842

Csernansky J G, Joshi S, Wang L, Haller J W, Gado M, Miller J P, Grenander U and Miller M I 1998 Hippocampal morphometry in schizophrenia by high dimensional brain mapping *Proceedings of the National Academy of Sciences* **95** 11406–11411

Davatzikos C 1996 Spatial normalization of 3d brain images using deformable models *Journal of Computer Assisted Tomography* **20** 656–665

Dice L R 1945 Measures of the amount of ecologic association between species *Ecology* **26** 297–302

Fowler J, Chappell R and Ritter M 2001 Is α/β for prostate tumors really low? *International Journal of Radiation Oncology*Biology*Physics* **50** 1021–1031

Fowler J F 1989 The linear quadratic formula and progress in fractionated radiotherapy *British Journal of Radiology* **62** 679–694

Goitein M 1985 Calculation of the uncertainty in the dose delivered during radiation therapy *Math. Phys.* **5** 608–612

Goitein M 1986 Causes and consequences of inhomogeneous dose distributions in radiation therapy *International Journal of Radiation Oncology*Biology*Physics* **12** 701–704

Goitein M and Busse J 1975 Immobilization error: Some theoretical considerations *Radiology* **117** 407–412

Grenander U and Miller M I 1998 Computational anatomy: An emerging discipline *Quarterly of Applied Mathematics* **56** 617–694

Happersett L, Mageras G S, Zelefsky M J, Burman C M, Leibel S A, Chui C, Fuks Z, Bull S, Ling C C and Kutcher G J 2003 A study of the effects of internal organ motion on dose escalation in conformal prostate treatments *Radiotherapy and Oncology* **66** 263–270

Joshi S, Lorenzen P, Gerig G and Bullitt E 2003 Structural and radiometric asymmetry in brain images *Medical Image Analysis* **7** 155–170

Joshi S, Miller M and Grenander U 1997 On the geometry and shape of brain sub-manifolds *Int. Journal of Pattern Recognition and AI* **11** 1317–1342

Joshi S C and Miller M I 2000 Landmark matching via large deformation diffeomorphisms *IEEE Transactions On Image Processing* **9** 1357–1370

Ketting C H, Austin-Seymour M, Kalet I, Unger J, Hummel S and Jacky J 1997 Consistency of three-dimensional planning target volumes across physicians and institutions *International Journal of Radiation Oncology*Biology*Physics* **37** 445–453

Kupelian P A, Reddy C A, Carlson T P, Altsman K A and Willoughby T R 2002 Preliminary observations on biochemical relapse-free survival rates after short-course intensity-modulated radiotherapy (70 gy at 2.5 gy/fraction) for localized prostate cancer *International Journal of Radiation Oncology*Biology*Physics* **53** 904–912

Leunens G, Menten J, Weltens C, Verstraete J and van der Scheuren E 1993 Quality assessment of medical decision making in radiation oncology: Variability in target volume delineation for brain tumours *Radiotherapy and Oncology* **29** 169–175

Lu W, Chen M, Olivera G H, Ruchala K J and Mackie T R 2004 Fast free-form deformable registration via calculus of variations *Physics in Medicine and Biology* **49** 3067–3087

Mageras G S, Joshi S, Davis B, Pevsner A, Hertanto A, Yorke E, Rosenzweig K and Ling C C Evaluation of an automated deformable matching method for quantifying lung tumor motion in respiration-correlated CT images In *ICCR* 2004

Miller M I, Trouve A and Younes L 2002 On the metrics and euler-lagrange equations of computational anatomy *Annual Review of Biomedical Engineering* **4** 375–405

Rasch C, Barillot I, Remeijer P, Touw A, van Herk M and Lebesque J V 1999 Definition of the prostate in CT and MRI: a multi-observer study *International Journal of Radiation Oncology*Biology*Physics* **43** 57–66

Schaly B, Kempe J A, Bauman G S, Battista J J and Dyk J V 2004 Tracking the dose distribution in radiation therapy by accounting for variable anatomy *Physics in Medicine and Biology* **49** 791–805

Thirion J P 1998 Image matching as a difusion process: an analogy with maxwell's demons *Medical Image Analysis* **2** 243–260

Thompson P M and Toga A W 2002 A framework for computational anatomy *Computing and Visualization in Science* **5** 13–34

Unkelbach J and Oelfke U 2004 Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions *Physics in Medicine and Biology* **49** 4005–4029 URL http://stacks.iop.org/0031-9155/49/4005

Valley J F and Mirimanoff R O 1993 Comparison of treatment techniques for lung cancer *Radiotherapy and Oncology* **28** 168–173

Wang H, Dong L, O'Daniel J, Mohan R, Garden A S, Ang K K, Kuban D A, Bonnen M, Chang J Y and Cheung R 2005 Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy *Physics in Medicine and Biology* **50** 2887–2905

Yan D, Jaffray D A and Wong J W 1999 A model to accumulate fractionated dose in a deforming organ *International Journal of Radiation Oncology*Biology*Physics* **44** 665–675

Yan D, Lockman D, Brabbins D, Tyburski L and Martinez A 2000 An off-line strategy for constructing a patient-specific planning target volume in adaptive treatment process for prostate cancer *International Journal of Radiation Oncology*Biology*Physics* **48** 289–302

Yan D, Vicini F, Wong J and Martinez A 1997 Adaptive radiation therapy *Physics in Medicine and Biology* **42** 123–132

Zijdenbos A P, Dawant B M, Margolin R A and Palmer A C 1994 Morphometric analysis of white matter lesions in mr images: Method and validation *IEEE Transactions on Medical Imaging* **13** 716–724

Zou K H, Warfield S K, Baharatha A, Tempany C, Kaus M R, Haker S J, Wells W M, Jolesz F A and Kikinis R 2004 Statistical validation of image segmentation quality based on a spatial overlap index *Academic Radiology* **11** 178–189